Mapping of ionospheric total electron content using global navigation satellite systems

Meggs, Robert W.

Award date:
2005

Awarding institution:
University of Bath

MAPPING OF IONOSPHERIC TOTAL ELECTRON CONTENT USING GLOBAL NAVIGATION SATELLITE SYSTEMS

Submitted by Robert W. Meggs
for the degree of
Doctor of Philosophy
of the University of Bath
2005

Robert W. Meggs

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary

Specification of total electron content (TEC) over a wide area is important for communications, surveillance and navigation. Established methods use either empirical models or two-dimensional mapping. A new technique for imaging the electron density in the ionosphere is explored. The technique produces four-dimensional images of the electron density and is known as MIDAS (Multi-instrument data analysis system). In this thesis MIDAS is used with dual-frequency differential-phase observations from the Global Positioning System (GPS) to produce images of the electron density over Europe.

A short simulation study shows that the MIDAS inversion produces more accurate slant TEC estimates than those produced by the two-dimensional mapping, for two locations in mainland Europe. A more extensive simulation study includes examples representing the ionosphere over an entire year at solar maximum. The results showed that the vertical TEC values were always more accurately produced using the MIDAS inversion than using the thin shell.

For the experimental case, three different approaches were adopted to assessing the images. In the first approach, the GPS data differential-code and differential-phase data were calibrated and the resulting slant TEC values were used to assess the images. In the second approach, a property of the receiver hardware biases, their stability, is exploited as an estimate of the relative accuracy of the MIDAS inversion compared to the two-dimensional mapping. In the third approach, a case study, the images are compared with measurements from independent instrumentation, the European Incoherent Scatter (EISCAT) radar. This showed that the main trough could be imaged by the MIDAS technique.

A simple graphical interface to MIDAS was developed for use by non-expert personnel. This demonstrates the use of the technique to the industrial sponsor for mapping the ionosphere over the European region.

The work of this thesis indicates that a significant improvement in wide-area TEC mapping can be achieved by using the new imaging method compared to previous techniques. The results have implications for improving the accuracy of satellite navigation systems and military communications and radar systems requiring ionospheric corrections.

Contents

List of Figures

List of Tables

List of abbreviations

AS	Anti Spoofing
BPSK	Binary Phase-Shift Keying
C/A	Course Acquisition
DGPS	Differential Global Positioning System
DOP	Dilution of Precision
ECEF	Earth Centred Earth Fixed
EGNOS	European Geostationary Navigation Overlay System
EISCAT	European Incoherent Scatter
EOF	Empirical Orthonormal Function
EUV	Extreme Ultra Violet
GPS	Global Positioning System
GUI	Graphical User Interface
HF	High Frequency (100 kHz to 30 MHz)
IFB	Inter-Frequency Bias
IRI	International Reference Ionosphere
ISR	Incoherent Scatter Radar
LEO	Low Earth Orbit
MIDAS	Multi-Instrument Data Analysis System
NNSS	Navy Navigational Satellite System
OCS	Operational Control Segment
PPS	Precise Positioning Service
PRN	Pseudo Random Number
Rx	Receiver
SA	Selective Availability
SOPAC	Scripps Orbit and Permanent Array Center
SPS	Standard Positioning Service
sTEC	slant Total Electron Content
TEC	Total Electron Content
TECu	Total Electron Content Unit ($= 10^{16}$ electrons m^{-2})
TOPEX	Typhoon Operational Experiment
Tx	Transmitter

UERE	user Equivalent Range Error
UHF	Ultra High Frequency (300 MHz to 1 GHz)
UT	Universal time
VHF	Very High Frequency (30 MHz to 300 MHz)
vTEC	vertical Total Electron Content
WAAS	Wide-Area Augmentation System
WADGPS	Wide-Area Differential GPS
WDC	World Data Centre

List of Symbols

b_i	=	measured or simulated total electron content (TEC) along ray path i
c	=	speed of light in free space
e	=	charge on an electron
f	=	radio-wave frequency
h	=	height
l	=	incremental distance along propagation path
m_e	=	mass of electron
n	=	refractive index of ionosphere
p	=	pseudorange from receiver to satellite
r_i	=	position of receiver i
s_j	=	position of satellite j
t	=	time
v	=	velocity of propagation
x_j	=	electron density in pixel j
z_i	=	residual between measured and computed TEC along ray path i
A_{ij}	=	length of ray path element i through pixel j
B_r	=	receiver inter-frequency bias
B_s	=	satellite inter-frequency bias
E	=	elevation angle of satellite
I	=	total electron content (TEC) through the ionosphere
L	=	carrier phase range in cycles
M	=	path length error due to multipath
N	=	electron density in the ionosphere
P	=	PRN code range in metres
R	=	radial distance from the centre of the Earth
S	=	excess path due to ionospheric delay
ϵ	=	dispersive component of phase range hardware delay
ε_0	=	permittivity of free space
λ	=	wavelength of transmitted signal in free space
μ	=	real part of ionospheric refractive index
ν	=	electron-neutral collision frequency

ρ = true geometric range from satellite to receiver
σ = standard deviation
ϕ = signal phase
χ = imaginary part of ionospheric refractive index
ω = angular frequency
Θ = geographic latitude
Λ = geographic longitude
τ = dispersive component of code range hardware delay

Matrices and vectors

$\mathbf{b}$ = vector of measured TEC
$\mathbf{c}$ = vector of changes in measured TEC
$\mathbf{r}$ = true geometric range vector from receiver to satellite
$\mathbf{s}$ = satellite position vector
$\mathbf{u}$ = user position vector
$\mathbf{x}$ = vector of electron density in each voxel
$\mathbf{y}$ = change in electron density
$\mathbf{z}$ = vector of residuals
$\mathbf{A}$ = matrix of path lengths through each voxel
$\mathbf{D}$ = change in ray path geometry
$\mathbf{G}$ = linear changes in coefficients of basis functions
$\mathbf{M}$ = relative contributions of basis functions
$\mathbf{X}$ = mapping matrix

Acknowledgements

I would like to acknowledge the support provided by the UK Engineering and Physical Sciences Research Council (EPSRC), and by my industrial sponsor BAE SYSTEMS. In particular, my thanks go to Mr. Glyn Wyman of BAE SYSTEMS for his interest and professional guidance. I would also like to thank the staff at the EISCAT Tromsø site for their practical help in operating the radar system, and the EISCAT group of the Rutherford-Appleton Laboratory for their help in analysing the radar data and their support for my visits to Tromsø. EISCAT is an International Association supported by Finland (SA), France (CNRS), the Federal Republic of Germany (MPG), Japan (NIPR), Norway (NFR), Sweden (NFR) and the United Kingdom (PPARC). I also acknowledge, with thanks, the use of the IRI model, and the SOPAC archive for GPS data.

On a personal note, I am indebted to my supervisor, Dr. Cathryn Mitchell, for the opportunity to pursue this research project. She has created an environment in which research is intellectually challenging, and yet pervaded by a sense of excitement and, dare I say, fun. Her patience, honesty and encouragement in helping me to rise to the challenge have been above and beyond the call of duty, and I have positively enjoyed working with her.

However, it would not have been possible for me to undertake this project without the support of my wife, Johannah. My thanks to her for all her love and support in all events that have led me to this stage in my life. I could not have done it without her.

I must also thank my friends and colleagues at the University of Bath. Space does not permit me to mention them all by name, but they know who they are. Thanks to them all for the many technical discussions and social happenings that we have shared together. Through the media of good company, good coffee, and the occasional beer they have led me to a better understanding of a wide variety of technical and scientific matters, and we've had a great time, too!

My chief regret is that my father, who saw me embark on this project, did not live to see me complete it. It was his love of, and enthusiasm for, engineering in all its forms that inspired me from a very early age, and so it is appropriate that I dedicate this work to him.

Chapter 1

Introduction

The ionosphere is the ionised region of the Earth's upper atmosphere, and is of fundamental importance to the understanding of radio propagation involving satellite communications. In particular, radio signals transmitted from Earth orbiting satellites to the ground are subject to an unknown time delay as they pass through the ionosphere. This unknown time-delay is directly proportional to the amount of ionisation along the signal path, known as the total electron content (TEC), and inversely proportional to the square of the signal frequency. Thus, for a given TEC value along a particular path, signals of higher frequencies will suffer smaller delays than those of lower frequencies. This allows TEC to be measured by using two ranging signals of different frequencies. If sufficient measurements of TEC are made at the same instant, it is possible to map the spatial distribution of the ionisation over a wide geographical area. The Global Positioning System (GPS) satellites transmit two phase-coherent ranging signals at separate frequencies, and hence they are a convenient source of dual-frequency data for TEC mapping.

An established ionospheric TEC mapping method, known as the "Thin Shell" method, represents the ionosphere by an infinitely thin shell at a fixed altitude (usually around 400 km). The main disadvantage of this method lies in the loss of data caused when the three-dimensional space of the ionosphere is mapped onto a two dimensional geometric space. A new method - the subject of this thesis - is an advanced tomographic algorithm that has been extended into four-dimensions.

1.1 An introduction to the Ionosphere

The ionosphere is the region of electrically conducting plasma in the Earth's atmosphere, ranging from approximately 50 km to over 1000 km in altitude. It is formed when extreme ultra violet (EUV) radiation and X-rays from the sun ionise the neutral atmosphere. The density at these altitudes is very low, and so electrons and ions produced by the ionisation process are free to move around in the plasma, before eventually recombining with ions to form neutral atoms. The distribution of the ionisation is highly variable, exhibiting both space and time dependencies.

In latitude, the ionosphere displays widely differing characteristics. The equatorial ionosphere is characterised by a fountain effect in which free electrons experience a strong upwards drift under the influence of the Earth's magnetic field. This drift, sometimes called the $\mathbf{E} \times \mathbf{B}$ drift, redistributes electrons over the equator to higher latitudes up to about 20° north and south. Strong horizontal electron density gradients are evident in this region as a result. At mid-latitudes the plasma is driven by neutral winds and electric fields, and follows the geomagnetic field lines. Consequently, the plasma tends to co-rotate with the Earth. Electron-density gradients tend to be shallower than those of the equatorial ionosphere. The morphology of the polar ionosphere is highly complex, being mainly convection driven. The Earth's magnetic field lines are open in this region, leading to coupling between the solar wind and the interplanetary magnetic field. Between the polar and the mid-latitude regions lies the auroral region, which is characterised by particle precipitation. Longitudinally distributed depletions in electron density known as 'troughs' are frequently evident in this region.

In the vertical domain, up to four peaks in electron concentration may form at distinct altitudes in the ionosphere. This is partly due to the fact that the neutral gas density decreases with height, so that there are fewer neutral atoms to participate in the ionisation process. On the other hand, the intensity of the solar EUV radiation increases with height. The depth to which the EUV radiation can penetrate depends on its wavelength, and different gas species are ionised more or less strongly by different wavelengths of EUV radiation. With the stronger radiation in the daytime, up to four distinct layers may form:

- The F2 region: Height about 210 km and above, with electron density peak typically in the range 250 - 400 km. The F2 region has the greatest electron density and is the most variable. It is composed mostly of atomic oxygen;
- The F1 region: Height about 140 - 210 km. It is composed of various molecular species;

- The E region: Height about 90 - 140 km. Typically has three different forms:
  - Normal: Due to solar soft X-rays;
  - Auroral: Has irregular structure caused mainly by particle precipitation;
  - Sporadic E: A thin layer of irregular occurrence.
- The D region: Height about 50 - 90 km. Formed by hard X-rays and Lyman α radiation acting on various molecular species.

When plotted as a function of height, the electron density forms a distinct profile, such as is shown in Figure 1.1. During the night time, only the F2 region remains, since there is no ionising radiation present and the recombination time is longer due to the lower density at higher altitudes. The upper limit of the F2 region is generally regarded to be the height at which hydrogen ions start to dominate. This region, known as the protonosphere, has very low density and may extend to altitudes in the order of 22,000 km (J.A. Klobuchar, in Parkinson *et al.* (1996)).

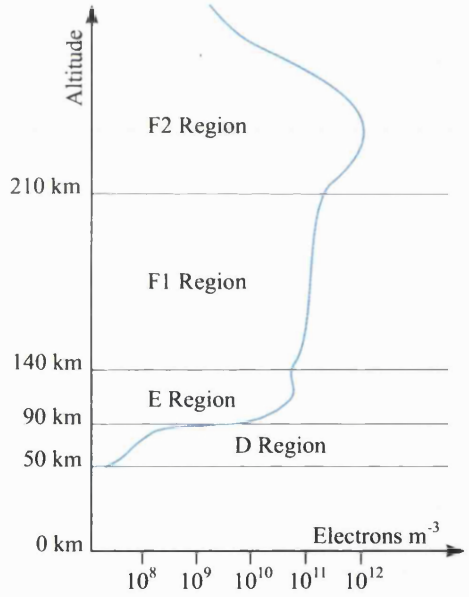


Figure 1.1: Typical form of daytime ionospheric vertical electron density profile.

The diurnal variation of the ionosphere follows a clearly defined pattern. At dawn, solar radiation starts to produce free electrons in the ionosphere. The F-region decreases in altitude during the morning, and may divide into two electron density peaks known as F1 and F2. The E and the D regions also start to form at lower altitudes. At mid-latitudes, the ionisation reaches a maximum early in the post-noon period. As the afternoon progresses, recombination and other electron-loss processes start to dominate as the intensity of the ionising radiation decreases. After local sunset, the D, E and F1

regions become insignificant, but the F2 region rises to higher altitudes and remains present throughout the night.

In the longer term, the ionosphere is subject to seasonal changes due to variations in the solar zenith angle. This would lead to the expectation of an overall increase in ionisation during the summer months. However, this is not necessarily the case, as there are also seasonal variations in the composition of the neutral atmosphere. In addition, the Sun undergoes a cyclic rise and fall in intensity, lasting, on average, approximately 11 years, although it can vary in length from 9 to 14 years. At solar maximum, the Sun emits a greater intensity of radiation and causes enhanced electron concentration in the ionosphere. Thus the long term temporal variability of the ionosphere is complex, and the reader is referred to McNamara (1991), Hargreaves (1992) and Davies (1990) for further information.

To radio signals, the ionosphere is a refractive medium, with the index of refraction being dependent upon the amount of ionisation present. Signals incident upon the ionosphere are subject to time delay, phase advance, refraction, reflection and rotation of the electric field vector (Faraday rotation) in inverse proportion to the square of their frequency. The ionosphere is therefore said to be dispersive, with these effects tending to be stronger at lower frequencies. For example, when signals in the frequency range 3 - 30 MHz (termed High-Frequency, HF) are obliquely incident upon the ionosphere, they are refracted and arrive back on the ground several hundred kilometres from where they originated. This effect was first demonstrated by Marconi in his 1901 experiment, in which he successfully transmitted a signal from Poldhu, UK, to St. Johns, Newfoundland. Since then the ionosphere has been widely exploited as the propagation medium for worldwide HF communications. Signals at Very High Frequency (VHF - over 30 MHz) and above are less strongly affected by the ionosphere. Above about 100 MHz, radio signals pass through the ionosphere with relatively little absorption. Further details on the ionospheric effects on radio signals can be found in Davies (1990).

1.2 Project Objectives

The aim of this work is to evaluate a new multi-dimensional inversion technique that uses GPS data to map the spatial distribution of ionisation in the ionosphere over a wide geographical area. The parameter of interest is the TEC, which is defined as the integrated value of the electron concentration along a column of unit cross section. The process of mapping the ionisation in the ionosphere is therefore known as TEC mapping.

The main objectives of this project are:

- To compare thin shell and advanced tomographic inversions by simulation;
- To assess the quality of the advanced tomographic inversion method experimentally;
- To verify the advanced tomographic inversion method using independent instrumentation;
- To develop a simple user interface to the advanced tomographic algorithm.

1.3 Thesis overview

The measurement of the ionosphere used this project is the total electron content (TEC), which is defined at the start of Chapter 2. TEC is easily measured by range measurements through the ionosphere on two phase-coherent signals of different frequencies. A convenient source of suitable signals is the GPS system. Accordingly, a presentation of the GPS system is then given in Chapter 2, starting with a description of space, control and user segments, followed by the signal generation system. A brief discussion on the typical positioning accuracy that can be expected from GPS is also presented, and the method of determining a user's position is outlined. Chapter 2 continues with descriptions of two independent instruments that are also used in this project. The first is the ionosonde, which is a ground-based vertically looking radar capable of measuring the maximum frequency that will be returned from the ionosphere. The second independent instrument is the European Incoherent Scatter Radar (EISCAT) radar system at Tromsø, Norway, from which independent data was obtained for verification purposes. An ionosonde that is co-located with the EISCAT radar was used to calibrate the radar data.

Methods of measuring and mapping TEC are discussed in Chapter 3. A general discussion on some applications of TEC mapping is given. Remote sensing techniques have also enabled the development of empirical models of the ionosphere, and a few of these are outlined. The TEC measurement is then developed, starting from the Appleton-Hartree equation. The thin shell method of mapping TEC is presented, and some of its shortcomings are discussed. An introduction to tomography is then presented, and its application to ionospheric imaging is discussed. A new 4-dimensional imaging method developed from 2-dimensional tomography, known as Multi Instrument Data Analysis System (MIDAS), is introduced, and the MIDAS algorithm is described. In the final section, a description of how the MIDAS algorithm is applied in this project.

A preliminary study based on a simulated ionosphere is described in Chapter 4. The purpose of this study was to compare the ionospheric determination of the full 4-dimensional MIDAS inversion with the established method based on a thin shell geometry. Using a minimal network of GPS receivers in Europe, reconstructions for three days were made using both methods. Two test locations in Europe were chosen to represent GPS single frequency users requiring an ionospheric position correction. It was found that the MIDAS inversion method consistently led to smaller position errors for the single-frequency user. This helped to provide the impetus for further investigation.

In Chapter 5, results from a long-term study using both simulated and real ionospheric measurement data are presented. A network of ground-based GPS receivers in Europe has been selected for both the simulated and the experimental cases. Twenty-five days were chosen from the year 2001 on the basis of at least one geomagnetically quiet and one geomagnetically active day per month. In the simulated case, a model ionosphere for each day was produced using the International Reference Ionosphere 1995 model (IRI-95), and inversions were made using the thin shell and the full MIDAS inversion. The resulting vertical TEC maps were compared with the vertical TEC distribution in the original model. Vertical TEC residuals for both methods were plotted as contour plots to show the spatial variation of the errors. Examples are presented for one of the 25 days. The errors were also plotted as the mean residual TEC for each hour of the 25-day data set. It was found that, in general, the vertical TEC errors in the thin shell cases were about three times as large as those in the full inversion. In the experimental case, the measured slant TEC was compared with the slant TEC through the full inversion, and error statistics are presented for all 25 days.

In Chapter 6, the experimental work of Chapter 5 is extended by exploiting the stability of the receiver hardware biases as an indirect measure of the TEC error. Three days in July 2000 were chosen, with geomagnetic conditions ranging from quiet to very disturbed. In the cases examined, it was found that the full inversion (MIDAS) approach gave an improvement over the thin shell for TEC mapping.

Images generated by independent instrumentation are compared with MIDAS-generated images in Chapter 7. The EISCAT radar at Tromsø, Norway, was operated in January 2002 to produce images of the winter-time ionosphere over northern Scandinavia. These were then compared with electron-densities generated by the MIDAS algorithm, using as input only GPS data from a network of receivers in Europe. A description of the EISCAT experiment is given, and images presented showing the southerly migration of the dayside trough. Importantly, a close correspondence in the trough location and shape can be observed between the radar images and the GPS images.

A design for a graphical interface for the MIDAS software is presented in Chapter 8.

This was requested by the industrial sponsor as a simple method of presenting ionospheric delay data to a non-expert user. A 7-band classification of the TEC at a given location in Europe, and at a specified time on a particular day was devised. The classification makes use of the statistics of TEC for the previous 30 days at that particular time of day. A simple interface informs the user if the TEC was anomalously low or high.

In Chapter 9, the work of the main results chapters is drawn together, and some general conclusions on the work as a whole are given. Possibilities for the future direction of the work are discussed.

Chapter 2

Measurement techniques and instrumentation

Abstract

TEC is the fundamental measurement of the amount of ionisation in the ionosphere. In the first section of this Chapter, TEC is defined as the integrated value of electron concentration along a path of unit cross section through the ionosphere. Descriptions of the instruments used in this project to measure TEC follow, beginning with the GPS system. The process of computing the user's position using measurements of the satellite to receiver path lengths is also outlined. This is followed by a description of the ionosonde, which is a simple radar that determines the height of the reflecting layers in the ionosphere. Finally, the European Incoherent Scatter (EISCAT) radar system is described in outline. Data from the EISCAT radar are used in this project for independent verification of the GPS images. The radar data were calibrated using data from a co-located ionosonde.

2.1 Total Electron Content (TEC)

Knowledge of the distribution of the ionisation can be used to determine the propagation characteristics of a signal path and mitigating the environmental effects of the ionosphere. An important parameter from which the amount of ionisation can be inferred is the total electron content (TEC), which is defined as follows. It is assumed that a transmitter is mounted on a satellite in orbit above the ionosphere, and a receiver

is situated on the Earth's surface. Then the TEC, I, is defined as the total number of electrons in a column of unit cross section from the transmitter to the receiver, i.e.,

$$I = \int_r^s N(R, \Theta, \Lambda).dl \quad (\text{ electrons m}^{-2}\ ) \tag{2.1}$$

where N is the electron concentration in electrons m^{-3}, R is the radial distance from the centre of the Earth, Θ is geographical latitude, Λ is geographical longitude, and l is the distance along the signal path from the satellite s to the receiver r. TEC may be expressed as the direct slant path between a transmitter and a receiver (sTEC), or an equivalent vertical TEC between a point on the Earth's surface and its zenith (vTEC). In either case, TEC is expressed in TEC units (TECu), where 1 TECu is equivalent to 10^{16} electrons m^{-2}. Observed values of TEC range from about 10^{16} electrons m^{-2} up to 10^{19} electrons m^{-2} (J. A. Klobuchar, in Parkinson *et al.* (1996)). It is important to note that single TEC measurements contain no information about the distribution of the electron concentration along the path.

In the past, TEC has been measured using the Faraday rotation of signals from geostationary satellites (see, for example Titheridge (1972)), and a database of TEC measurements by Faraday rotation is now in existence (Davies and Hartmann, 1997). However, a problem associated with TEC measurements made by Faraday rotation is that neither the initial direction of the electric field, nor the number of cycles of rotation are known (Davies, 1990). It is now more usual to infer TEC from the dispersive effect of the ionosphere on dual-frequency signals.

The Navy Navigation Satellite System (NNSS) and the GPS satellites offer suitable signals for this purpose. The NNSS satellites are in polar orbits at an altitude of about 1100 km, and transmit two phase-coherent signals on 150 MHz and 400 MHz. The GPS satellites orbit at altitudes of about 20200 km, and transmit two phase-coherent signals at L-band frequencies of 1575.42 MHz and 1227.60 MHz, known respectively as L1 and L2. The carrier frequencies used by GPS system are sufficiently high for the effects of ionospheric absorption and the Earth's magnetic field to be ignored. A detailed description of the GPS system is given in the next section.

2.2 The Global Positioning System (GPS)

2.2.1 An introduction to GPS

The Global Positioning System (GPS) is a satellite-based navigation and positioning service providing users with a worldwide high accuracy positioning capability in three dimensions, together with time and velocity determination. It is operated by the US Department of Defense (DoD), and became fully operational in July 1995. Two levels of service are offered: the Precise Positioning Service (PPS), and the Standard Positioning Service (SPS). The PPS is encrypted, and is primarily intended for US armed services and government agencies, whilst the SPS is available, at reduced accuracy, to civil users worldwide. Initially, the accuracy of the SPS was further degraded by a policy known as Selective Availability (SA). However, a massive increase in civilian demand for GPS products led the US government to switch off SA in May 2000.

The GPS system consists of three main segments: the space segment; the Operational Control Segment (OCS); and the user segment. In the space segment, there are at least 24 satellites, similar to the one illustrated in Figure 2.1, arranged in six orbits. Four operational satellites occupy each orbit, and additional satellites fly as spares. The

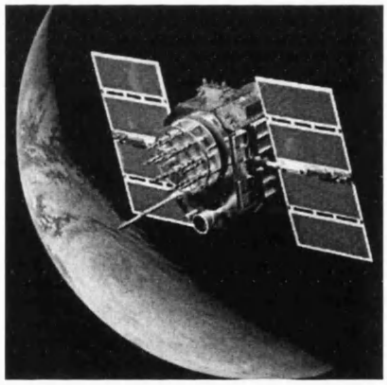


Figure 2.1: A GPS Block II satellite (source: www.fas.org)

orbits are inclined at 55° to the equatorial plane, and have a radius of approximately 26,600 km, giving an orbital period of one-half sidereal day (11h 58min). This orbital arrangement seeks to ensure that there are at least four satellites in view from any point on the Earth's surface. The satellites continuously broadcast time and ranging signals on two separate carrier frequencies of 1575.42 MHz and 1227.60 MHz, known respectively as L1 and L2.

The satellites are continuously monitored by the ground-based OCS, which manages

the tracking, telemetry and control functions. These include maintenance of station keeping, updating the ephemeris and almanacs, and monitoring system health. Since the satellite clocks are all free running, an important function the OCS is to maintain the synchronism of the satellite clocks to GPS system time, which is referenced to UTC maintained by the US Naval Observatory.

The user segment is the community of users, both military and civil, equipped with GPS receivers that receive and process the L-band signals transmitted by the satellites. There are two types of receiver: dual frequency and single frequency. Dual-frequency receivers are able to track both the L1 and the L2 frequencies simultaneously, and can be further categorised as military or commercial. Military receivers have access to a classified algorithm that enables them to recover the encrypted PPS ranging codes with the high dynamics necessary for applications such as aircraft navigation. Commercially available dual-frequency receivers do not have this ability, and so have to resort to codeless techniques to recover the PPS ranging codes. They are nevertheless capable of very high accuracy, but at the expense of protracted measurement times and high capital cost. Single-frequency receivers track the L1 frequency only, and recover the SPS ranging codes. They are considerably less expensive than dual-frequency ones, but are not capable of achieving the precision of dual-frequency receivers.

2.2.2 GPS signal generation

The GPS signals are derived from a master frequency reference of 10.23 MHz, with a stability of about 1 part in 10^{13}. Figure 2.2 shows the principal signal paths in a GPS Block II satellite. The master clock signal is multiplied up to produce two L-band carrier frequencies of of 1575.42 MHz and 1227.6 MHz. The navigation message is modulated onto both carriers with a data rate of 50 Hz, using binary phase shift keying (BPSK). The resulting signals are then spread using two different pseudo-random noise (PRN) codes, prior to final radio frequency (RF) amplification and transmission.

The two PRN codes are digital sequences that exhibit some of the characteristics of gaussian white noise, but are entirely deterministic. They are known as the 'Precise' (or 'P') code and the 'Coarse Acquisition' (or 'C/A') code. The P-code has an effective wavelength of 30 m, and simultaneously modulates the L1 and the L2 carriers. The PPS is delivered through the P-code. The public C/A-code has an effective wavelength of 300 m and modulates the L1 frequency only, in quadrature with the P-code modulation. The SPS is delivered through the C/A-code. Each satellite is identified by a unique PRN code, which it transmits together with a time-stamped ranging signal and a navigation message containing data about the satellite's position, clock offset

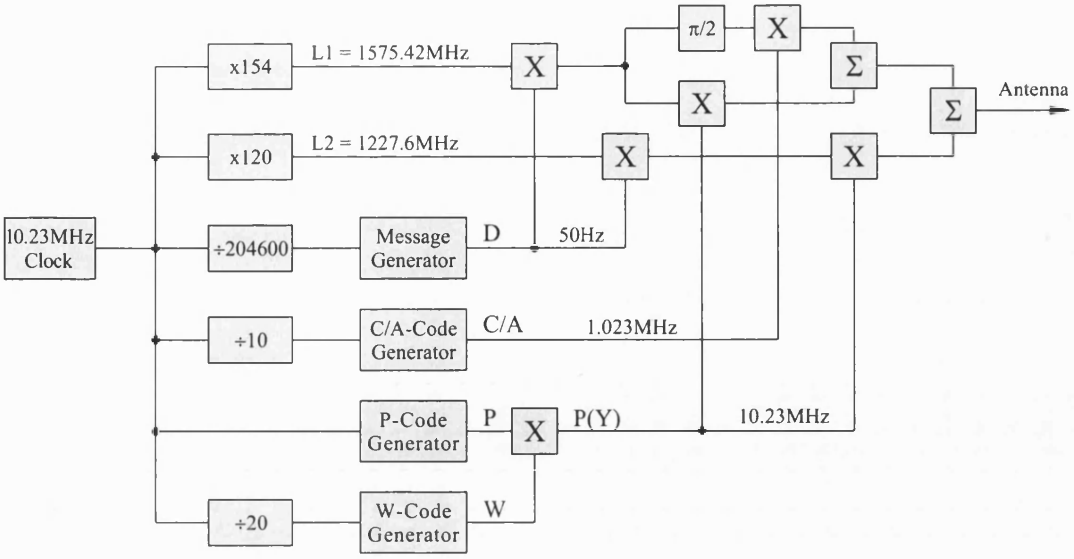


Figure 2.2: A simplified block diagram showing the principal signal paths in a GPS Block II satellite.

and health. A policy known as 'anti-spoofing' (AS) is implemented on the P code by modulating it with an encrypted code known as the 'W' code, which is believed to have a chipping rate of 511.5 kHz. The resulting code, known as 'P(Y)', is then modulated onto the two carrier frequencies. The intention of AS is to prevent the malicious 'spoofing' of the GPS ranging signals to mislead legitimate users of the PPS.

2.2.3 Positioning accuracy

The precision of the GPS navigation solution is limited by timing errors, ephemeris prediction and propagation errors[1]. These error terms are often lumped together and expressed as a factor, called the user equivalent range error (UERE). Typical values are listed in Table 2.1, from which it can be seen that the main limitation to the positioning accuracy of civil (single-frequency) GPS is due to the unknown time delay through the ionosphere.

The majority of the ionospheric delay affecting the L-band GPS signals during the daytime is due to the F2 region (between 250 and 400 km in altitude). When present, the F1 region (about 200 km) contributes up to 10% of the GPS signal delay. The E

[1]For an unambiguous position fix range measurements to a minimum of four satellites are required. Hence there is also uncertainty associated with the satellite-receiver positioning. This geometric uncertainty is known as dilution of precision (DOP). Since DOP is determined independently of the propagation characteristics of the signal path, it is outside the scope of this project, but further information can be found in Kaplan (1996).

Table 2.1: Typical error terms for PPS (P-code) and SPS (C/A-code)

Segment	Error source	UERE contribution (metres)	
		P-code	C/A-code
Space	Frequency standard stability	6.5	6.5
	Satellite acceleration uncertainty	2.0	2.0
	Other (thermal radiation, etc.)	1.0	1.0
Control	Ephemeris prediction	8.2	8.2
	Other (thrusters performance, etc.)	1.8	1.8
User	Ionospheric delay	4.5	9.8 - 19.6
	Tropospheric delay	3.9	3.9
	Receiver noise and resolution	2.9	2.9
	Multipath	2.4	2.1
	Other (inter-channel bias, etc.)	1.0	1.0
Total rms system UERE (metres)		13.0	15.7 - 23.1

Source: Navstar GPS User Equipment Introduction, public release version, September 1996, published by the Navstar GPS Joint Program Office. PDF file downloaded from http://gps.losangeles.af.mil/, 5th December 2003.

region (about 100 km) has minimal effects on GPS signal delay, although the irregular structure associated with auroral E-region activity can cause minor scintillation effects on GPS signals. The D region (about 80 km) has no measurable effect on GPS signals. At night, only the F2 region remains, and TEC levels fall to about 3 TEC units (equivalent to about 1 ns delay) in middle latitudes. The protonosphere is thought to contribute up to 10% of the daytime delay and up to 50% of the night time delay, although further study of the protonosphere is outside the scope of this project. Large vertical TEC values in excess of 220 TECu were recorded by Yin, *et al.* (2004) over the south-eastern United States during the storm of 15th July 2000 . This is approximately equivalent to an excess path length of 35 metres on the GPS L1 frequency (1575.42 MHz).

Dual-frequency GPS receivers can estimate the ionospheric delay directly by comparing the P-code delays on the L1 and L2 frequencies. Since single-frequency receivers track only the C/A-code on the L1 frequency, they cannot measure the ionospheric delay directly. Usually, a mathematical algorithm programmed into the receiver's firmware is used to estimate the ionospheric delay. A commonly used algorithm for this purpose is due to Klobuchar (1987). Alternatively, the ionospheric delay can be estimated by external means and a correction applied to the position solution through a data link to the receiver.

One approach to the problem of estimating the ionospheric delay uses one or more

reference GPS receivers that are fixed in precisely surveyed positions. The reference receivers compute the error by differencing their known and measured positions, and broadcast correction messages to nearby users on a separate radio frequency. Such an approach is known as Differential GPS (DGPS), and can provide improvements in positioning accuracy, provided that users are not more than a few tens of kilometres from the nearest reference receiver. However, a serious disadvantage of DGPS is that the positioning accuracy degrades significantly as the distance between the users and the reference receiver increases, making DGPS suitable only for small coverage areas such as harbour approaches.

An extension to DGPS uses data from a widely distributed network of reference receivers to map vertical TEC over a wide geographical area. Since TEC is proportional to time delay, mapping TEC in this way is equivalent to mapping the distribution of the ionospheric time delay for single frequency GPS users. Such an approach is currently operational in the United States as a wide-area correction system for satellite-based civil aircraft navigation. A similar system is under development in Europe, and other regions are considering the development of similar systems. The vision is that satellite-based navigation would eventually replace the existing global network of ground-based HF and VHF based radio navigation aids for all phases of flight. The reduced reliance on ground-based equipment opens up the possibility of making blind landing approaches at any airfield.

2.2.4 User position determination by GPS

A GPS receiver computes the user's position by measuring the lengths of the paths from at least four satellites to the receiver's antenna. The length of a satellite to receiver path is estimated by measuring the propagation time of the L-band signal. The navigation message contains the current position of the satellite and a time stamp indicating the time at which the signal left the satellite. The receiver internally generates copies of all the satellites PRN codes and cross-correlates each of these with the PRN codes received from the satellites in view. By differencing the time stamp and the receiver's internal time the propagation time of the signal, and hence the path length, can be computed.

Assuming perfect time synchronisation between the receiver clock, satellite clocks and the GPS system time a minimum of three observations to separate satellites, giving a system of equations in x, y and z, is necessary for an unambiguous position fix. Since receivers generally use a simple quartz clock, there are significant offsets of the receiver clock from GPS system time. The satellite clocks, being free running, also drift from GPS system time, but these are monitored and corrected by the OCS and

can be assumed to be an accurate representation of GPS time. An observation to a fourth satellite is therefore necessary, leading to a system of four linear equations in four unknowns (latitude, longitude, height, and receiver clock offset from system time). The solution to these equations is known as the navigation solution.

Ignoring all error terms except the satellite and receiver clock offsets, the navigation solution can be developed as follows (Kaplan, 1996). The positions of the satellite and the user's receiver are defined with reference to an Earth-centred Earth-fixed (ECEF) co-ordinate system, centred at the centre of mass of the Earth. The $+x$-axis is coincident with the equatorial plane, and points in the direction of 0° longitude; the $+z$-axis coincides with the spin axis of the Earth, and points in the direction of the North Pole; and the $+y$-axis completes the right-hand reference frame by pointing in the direction of 90° east.

Referring to Figure 2.3, let $\mathbf{u}$ be the vector representing the user's position, which is unknown, and $\mathbf{s}$ be the vector representing the satellite's position. Assume that $\mathbf{s}$ has been computed from the ephemeris data transmitted by the satellite. Then the vector

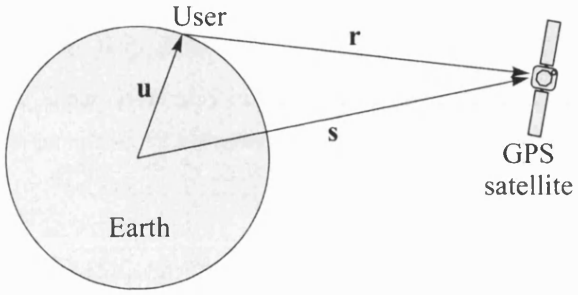


Figure 2.3: Range, user and satellite position vectors.

$\mathbf{r}$ is the true geometric range vector from the satellite to the user's receiver, and can be written

$$\mathbf{r} = \mathbf{s} - \mathbf{u} \tag{2.2}$$

The geometric range is the magnitude of $\mathbf{r}$, i.e. $\|\mathbf{r}\| = \|\mathbf{s} - \mathbf{u}\|$. Let ρ denote the true geometric range, so that

$$\rho = \sqrt{(x_s - x_u)^2 + (y_s - y_u)^2 + (z_s - z_u)^2} \tag{2.3}$$

In the absence of clock-offset errors, the true geometric range ρ from receiver r to satellite s can also be expressed as

$$\rho = c\,(T_r - T_s) \tag{2.4}$$

where c is the speed of light in a vacuum, T_r is the time at which the signal reached

the receiver, T_s is the time at which the signal left the satellite. However, since the satellite and receiver clocks are free running, they are offset from GPS system time by different amounts. Suppose that the satellite clock is offset from GPS time by δt_s and the receiver clock is offset by δt_r. Then the pseudorange, p, is the true geometric range that has been biased by clock, propagation and hardware errors, and is given by

$$\begin{aligned} p &= c[(T_r + \delta t_r) - (T_s + \delta t_s)] \\ &= c(T_r - T_s) + c(\delta t_r - \delta t_s) \\ &= \rho + c(\delta t_r - \delta t_s) \end{aligned} \tag{2.5}$$

The value of the satellite clock offset, δt_s, can be assumed to be correct since it is determined by the OCS, and corrections are included in the navigation message. Thus ignoring δt_s, Equation 2.5 above becomes

$$p = \rho + c\delta t_r \tag{2.6}$$

Hence the pseudorange from satellite s to receiver r can be written

$$p_{sr} = \sqrt{(x_s - x_r)^2 + (y_s - y_r)^2 + (z_s - z_r)^2} + c\delta t_r \tag{2.7}$$

To determine the user's position and clock offset, pseudorange measurements to four separate satellites are required. This gives a system of four equations,

$$\begin{aligned} p_{s_1r} &= \sqrt{(x_{s_1} - x_r)^2 + (y_{s_1} - y_r)^2 + (z_{s_1} - z_r)^2} + c\delta t_r \\ p_{s_2r} &= \sqrt{(x_{s_2} - x_r)^2 + (y_{s_2} - y_r)^2 + (z_{s_2} - z_r)^2} + c\delta t_r \\ p_{s_3r} &= \sqrt{(x_{s_3} - x_r)^2 + (y_{s_3} - y_r)^2 + (z_{s_3} - z_r)^2} + c\delta t_r \\ p_{s_4r} &= \sqrt{(x_{s_4} - x_r)^2 + (y_{s_4} - y_r)^2 + (z_{s_4} - z_r)^2} + c\delta t_r \end{aligned} \tag{2.8}$$

which must be solved for x_r, y_r, z_r and δt_r. Hofmann-Wellenhof *et al.* (2001) and Kaplan (1996) describe approaches for solving these that use analytical methods, iterative techniques based on linearisation, or Kalman filtering.

2.3 Ionosondes

An HF signal propagating through the ionosphere encounters layers of ionisation, which act like a mirror and reflects the signal back down to the ground. The height of the reflecting layer is termed the "virtual height", h', to convey the fact that the apparent reflection is really a refraction through layers at some height h_r, the real height (see Figure 2.4). The time taken for the signal to make the round trip depends on the virtual

height of the reflecting layer. For a signal at vertical incidence this can be found from the relation $h' = ct/2$. Higher frequencies penetrate further into the ionosphere before

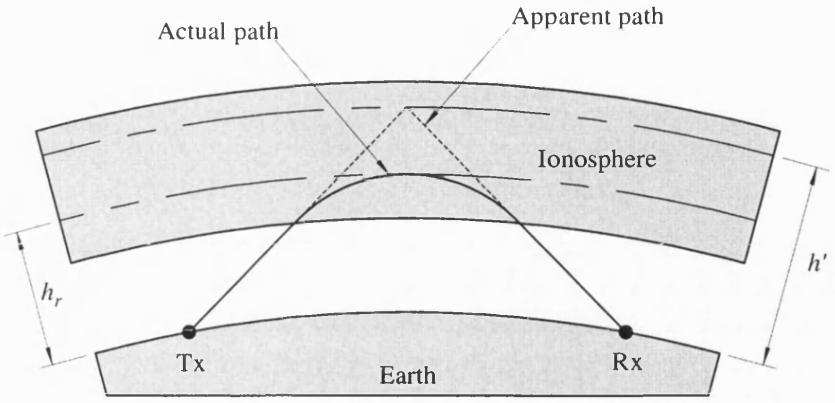


Figure 2.4: Definitions of virtual height (h') and real height (h_r) of a reflecting layer for an HF signal transmitted at oblique incidence.

being reflected. If the frequency of the signal is increased, a "critical frequency" f_c, is reached above which the signal passes through the ionosphere and out into space.

The ionosonde is one of the simplest instruments for direct measurement of the ionosphere. It measures the virtual height of the reflecting layers by transmitting pulses of radio energy vertically upwards and measuring the time t taken for the pulse to return. Successive pulses are transmitted at increasing frequencies. If the frequency is plotted against virtual height, the heights of the reflecting layers and the maximum frequencies associated with them can be deduced. Such a graph is called an ionogram, an example of which is shown in Figure 2.5. The maximum electron density, N_m, in a reflecting layer can then be found using the approximation

$$f_c \text{ (MHz)} \approx 9 \times 10^{-6}\sqrt{N_m} \qquad [\text{ electrons m}^{-3}\text{ }] \tag{2.9}$$

A similar technique can be used to observe the topside ionosphere from satellites orbiting above the ionosphere.

Ionosondes have a long and venerable history. In his Nobel lecture of 12$^{\text{th}}$ December 1947, Sir Edward Appleton describes an experiment that he carried out in 1924 using the BBC transmitter at Bournemouth and a receiver at Oxford to estimate the height of a reflecting layer. Experiments of this type were continued during the 1920s, notably by Breit and Tuve, and it was during this period that the modern ionosonde was born. Thus some ionosonde sites have been recording data since the 1930s.

Despite the wealth of ionosonde data available, ionosondes suffer from two major disad-

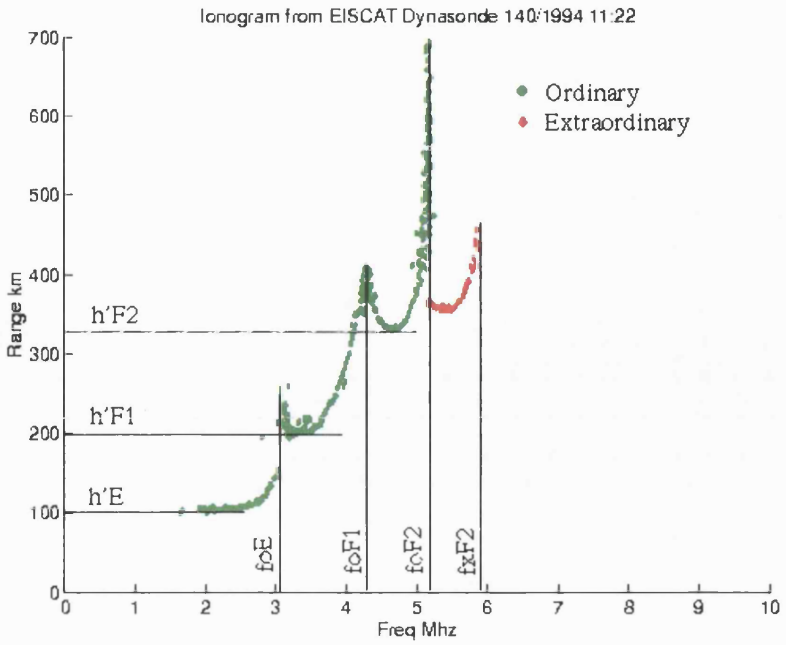


Figure 2.5: An example of a simple ionogram.
(source: http://www.wdc.rl.ac.uk/ionosondes/ionogram_interpretation.html)

vantages: they can only monitor the local ionosphere directly above their geographical position; and they can only observe up to the height of the F2 region peak density (the "bottomside"). They therefore do not give any direct information on the structure of the topside ionosphere, although some researchers have reported using ionosondes to infer topside ionisation, and hence TEC (see, for example, Reinisch *et al.* (2001)).

2.4 Incoherent scatter radar

Incoherent Scatter Radar (ISR) is a ground-based remote sensing technique for observing the Earth's near-space environment. It has developed into a powerful tool capable of measuring many of the properties of the ionosphere and neutral atmosphere. It offers two main advantages over the ionosonde (see section 2.3): it can observe the ionosphere above and below the electron density peak simultaneously, and it achieves better spatial resolution due to the narrow beam produced by the antenna, which is large relative to the radio wavelength (Hargreaves, 1992).

In common with any radar system, ISR emits a pulse of radio energy of known frequency

and receives echoes from targets in the path of the pulse. The difference between surveillance radar and ISR is that where the former detects a discrete target such as an aircraft, the latter receives returns from ion-acoustic waves of free electrons in the atmosphere. Consequently, the returns are very weak. This means that ISR requires large antennas, high transmitter power, highly sensitive receivers and sophisticated signal processing.

The radar returns are due to the motion of free electrons, and consist of spectral lines offset from the transmit frequency by an amount proportional to the electron velocity along the line of sight of the observation. The resulting data can be displayed as the ion-acoustic spectra shown in Figure 2.6.

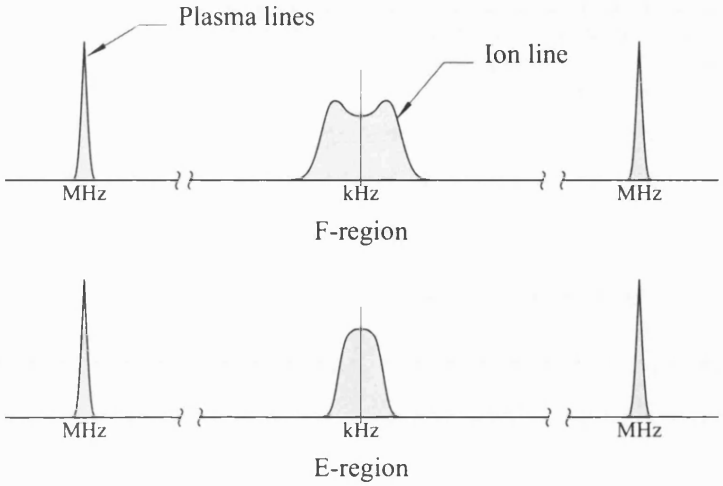


Figure 2.6: Ion-acoustic spectra from E and F regions.

The F-region return is due to scattering from random ion-acoustic waves, since the density of the plasma is relatively low. The up-going and downward-going ion-acoustic waves appear as a double peak in the ion line. In the E-region, the neutral gas density is much greater and so reflecting structures are more closely packed and ion-acoustic waves are unable to propagate. Consequently, the radar is unable to resolve individual structures and the ion line exhibits a single peak (Rishbeth and Williams, 1985). Figure 2.7 summarises some of the information that can be measured by incoherent scatter. Further background on incoherent scatter techniques is presented by Beynon and Williams (1978), and a description of the EISCAT radar and its early results is given by Rishbeth and Williams (1985).

Electron density data from the EISCAT radar has been used in this project for independent verification purposes, as described in Chapter 7. The EISCAT radar is situated at three sites in northern Scandinavia:

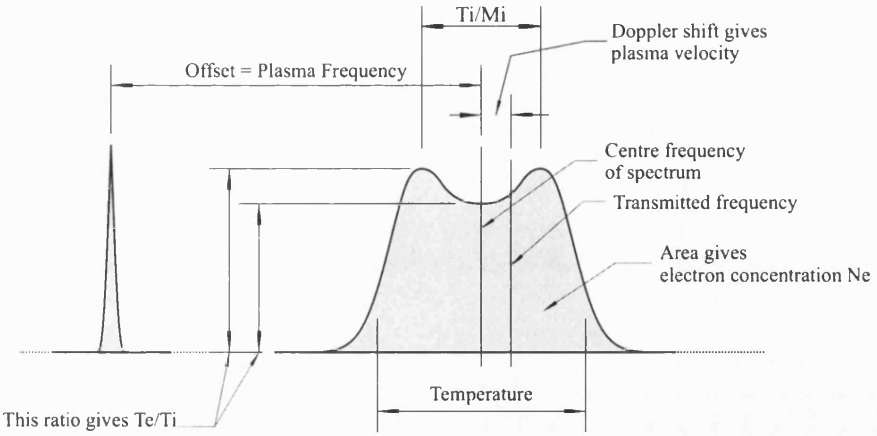


Figure 2.7: Parameters that can be measured by incoherent scatter radar.

- Tromsø, Norway: 69.58°N, 19.22°E.
- Kiruna, Sweden: 67.86°N, 20.43°E.
- Sodankylä, Finland: 67.36°N, 26.62°E.

The two transmitters are both located at the Tromsø site, which is illustrated in Figures 2.8 and 2.9. The VHF transmitter operates at about 224 MHz, and its two klystrons are capable of developing peak powers in the order of 3 MW. The VHF antenna is a section of a parabolic cylinder measuring 120 metres by 40 metres overall, and is steerable in the geographic meridian from 30° North elevation to 60° South elevation. The UHF transmitter is centred on a frequency of about 930 MHz, with a bandwidth of about 8 MHz. The peak operating power of the UHF system is in the order of 1.3 MW. A 32 metre diameter fully steerable dish antenna of the parabolic cassegrain type is provided. Receivers for both systems are dual-conversion types linked to a sophisticated signal processing system. The Kiruna and Sodankylä sites have no transmit capability, but are able to operate in tristatic mode with the receivers at Tromsø. Receiver gating in tristatic mode is controlled using GPS timing signals. A dynasonde is also situated at the Tromsø site. This is frequently used to calibrate the electron density data obtained from the radar. In addition to the three sites mentioned above, there is also a newer facility operational at Longyearbyen on Svalbard.

Figure 2.8: The EISCAT transmitter site, Tromsø, Norway, showing the main transmitter hall and the UHF antenna.

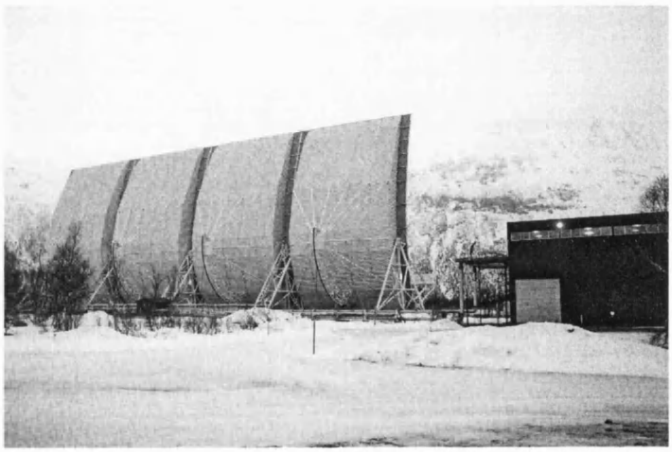


Figure 2.9: The EISCAT VHF radar antenna at Tromsø.

2.5 Summary

In this Chapter, TEC has been defined as the integrated value of electron concentration along a path of unit cross section through the ionosphere. Three different ionospheric measurement techniques and their instrumentation were then described. The first was the Global Positioning System (GPS), which is a satellite based navigation and timing system owned and operated by the United States government but available for public use. An overview of GPS was given, which was followed by a description of how the user's position is determined. The second instrument to be described was the ionosonde, which is a simple radar that determines the height of the reflecting layers in the ionosphere. Lastly, a brief description of the European Incoherent Scatter (EISCAT) radar system was given.

Chapter 3

Measurement and mapping of Total Electron Content

Abstract

In this Chapter, an introduction to TEC measurement and mapping is presented. The concept of TEC mapping is introduced in the first section, and some of its applications are discussed. These include: aircraft navigation; design and analysis of communications systems; and scientific research into the properties of the ionosphere. Ionospheric modelling is also introduced, and examples of some common models are presented. The development of the TEC measurement from dual-frequency GPS ranging measurements is given. Dual-frequency code-range TEC measurements are absolute, but noisy, whereas dual-frequency carrier-phase measurements possess low noise but are offset by an unknown integer-cycle ambiguity. The calibration of carrier-phase measurements to code-range measurements to produce a TEC measurement that is both smooth and absolute, is discussed. Two methods of mapping TEC over a wide geographical area are then introduced. The first is a well established method that represents the ionosphere as a thin shell. The second is a novel approach developed from 2-dimensional tomography, capable of producing movies showing the time evolution of the vertical TEC over a wide area. In the final section, a general description is given of the procedure adopted in this project to construct maps of ionospheric TEC using the two methods. The process of computing and removing the receiver and satellite inter-frequency biases is also explained.

3.1 Ionospheric observation and TEC measurement

In this section, the concept of, and approaches to, TEC mapping are introduced. Some applications of TEC mapping are suggested. In particular, the role of TEC mapping in computing ionospheric delay corrections for single-frequency GPS aircraft navigation is discussed, and a brief description is given of some efforts currently being undertaken to implement this. Other applications discussed include the use of TEC mapping in the design and analysis of radio communication systems, and in pure scientific research. Finally, the role of analytic models is discussed, and examples are presented.

3.1.1 TEC mapping and its applications

In Chapter 2, TEC was defined as the integrated value of electron concentration along a path of unit cross section between a satellite and a receiver. It is easily measured using differential ranging measurements on two separate frequencies, and permits the magnitude of the ionospheric signal delay at a single frequency to be estimated. The GPS satellites provide a convenient source of phase-coherent signals on two separate carrier frequencies that can be used to measure TEC. If a sufficient number of slant TEC measurements are made from a widely distributed network ground-based GPS receivers to the satellites in view, a map showing the horizontal distribution of the vertical TEC can be plotted. The ionospheric delay on paths between arbitrary points on (or near) the Earth's surface and points in space could then be computed.

In this project, two approaches to TEC mapping by direct measurement are compared. The first is the widely-used "Thin Shell" algorithm, described by Lanyi and Roth (1988), in which the ionosphere is represented by an infinitely thin shell at a fixed altitude near the F2 electron-density peak. The thin shell algorithm maps slant TEC measurements to the equivalent vertical TEC at the point of intersection of the measurement with the shell, using a geometric function of the elevation angle. Slant TEC on arbitrary paths are then retrieved by applying the inverse of the mapping function. The second approach is a new four-dimensional algorithm described by Mitchell and Spencer (2003). Here, the distribution of electron concentration along the paths of the slant measurements is computed from an inversion of linear equations. The resulting map is a three-dimensional image with a time dependency. Slant TEC on arbitrary paths are retrieved by integrating through the inversion. Furthermore, the time dependency permits movies to be made showing the time evolution of the TEC distribution. A fuller description of these algorithms is presented later in this chapter.

TEC mapping is currently being applied to the problem of estimating ionospheric delay

corrections for GPS navigation systems. Recall from Chapter 2 that the ionospheric delay constitutes the largest error term in the GPS navigation solution (see Table 2.1). Dual-frequency GPS receivers can eliminate directly the ionospheric delay, by making differential code or carrier phase measurements. On the other hand, single-frequency receivers rely on the Klobuchar model (or similar) to estimate the ionospheric delay. Under quiet geomagnetic conditions, the Klobuchar model is able to correct for, at best, about 50% of the ionospheric delay. The correction worsens significantly with the onset of ionospheric storm conditions (Klobuchar, 1987). Furthermore, there is no indication that the correction has deteriorated, since models can only work with 'average' conditions. However, wide-area TEC maps compiled from dual-frequency GPS measurements can be used to compute corrections for the ionospheric delay, which are subsequently broadcast to the users of single-frequency receivers. Thus mapping TEC over a wide geographical area can be regarded as equivalent to mapping the GPS signal delay through the ionosphere.

This principle has become known as Wide-Area Differential GPS (WADGPS). In the United States, the Wide Area Augmentation System (WAAS), described by El-Arini *et al.* (2001), is now operational. WAAS uses a thin shell algorithm to compute (in real time) ionospheric delay corrections for single-frequency GPS receivers. The WAAS correction message is uplinked to geostationary communications satellites and broadcast to civil aircraft operating in airspace over the continental United States. A further component of the WAAS correction message is an integrity indicator, driven by an ionospheric storm detector, that will inform users if the broadcast correction has become unreliable. A number of manufacturers are now producing WAAS-capable GPS receivers. A similar system, the European Geostationary Navigation Overlay System (EGNOS), is under development. It is intended to operate in European airspace and will use signals from the proposed European space-based navigation satellite network Galileo as well as GPS. There are also similar systems envisaged for regions as diverse as Brazil, India, Japan and South East Asia.

TEC mapping has applications in the design and analysis of terrestrial HF radio communications and direction finding. HF radio links operate in the 3 - 30 MHz range, and have advantages over systems operating at higher frequencies. For example, terminal equipment can be relatively inexpensive, and it can operate at low power levels, whilst still retaining sufficient bandwidth for voice and low rate data communications. Radar operating at HF has the ability to 'see' over the horizon. However, HF radio suffers some disadvantages. For example, the propagation conditions can be very variable, requiring frequent changes in operating frequency. Also, a large number of propagation paths are possible, leading to interference between the signals arriving by different paths and time delays (Davies, 1990). TEC mapping can help in identifying propagation paths, determine optimum operating frequencies and identify target locations for

military over the horizon radar systems.

The international scientific community has been keen to exploit transionospheric signals for remote sensing of the near-space plasma. Signals from NNSS and GPS have been employed in TEC mapping techniques to study space weather and its effects on the Earth's environment. TEC measurements have also been used to correct single-frequency altimeter measurements for ionospheric delay in the TOPEX/Poseidon missions (Schreiner *et al.*, 1997). TOPEX/Poseidon is a joint US/French satellite that monitors sea surface heights in the study of ocean variability and climate change. Early missions used a dual-frequency radar altimeter, which is able to correct for ionospheric delay. However, some later missions have used a single-frequency altimeter, the measurements of which are biased by the ionospheric delay. The purely scientific case for TEC mapping has thus been vindicated by the deeper understanding of the processes within, and morphology of, the ionosphere.

3.1.2 Ionospheric models

Current knowledge of the ionosphere has been assembled from a combination of ground based remote sensing techniques and satellite observations, made over many years. Yunck *et al.* (2000) present a good general review of the use of GPS for remote sensing of the ionosphere. Observations from remote sensing techniques have helped to develop empirical models that can be used to map ionospheric TEC without the use of direct measurements. Davies (1990) presents some simple models that approximate an ionospheric layer by an analytical expression. For example, the following model gives an estimate of the electron density N at a height h using a linear approximation of the electron density gradient α and the height h_0 of the base of the layer:

$$N = \alpha\,(h - h_0) \tag{3.1}$$

Another model described by Davies (1990) is the Chapman layer:

$$N = N_0\ \exp\{1 - z - \sec(\chi)\exp(-z)\} \quad \text{and} \quad z = \frac{(h - h_0)}{H} \tag{3.2}$$

where N_0 is the peak electron density at height h_0, H is the scale height, and χ is the solar zenith angle.

More complex models make use of peak density measurements combined with vertical electron density profiles to produce an estimate of the electron density at a particular height. An example of a widely used model is the International Reference Ionosphere

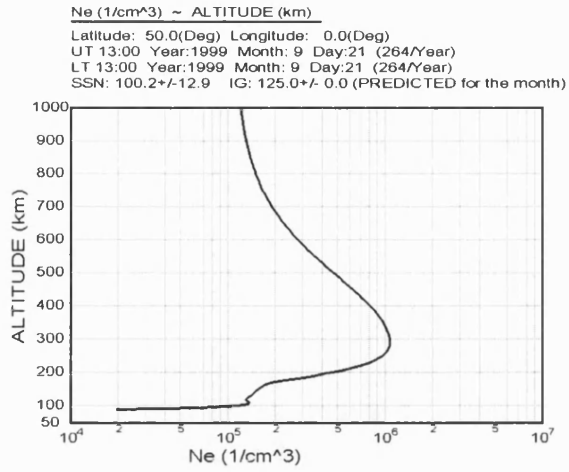


Figure 3.1: Example of an electron density profile computed using the IRI95 model.

(IRI - see Bilitza (1990)), which has been used extensively in this project to create a model ionosphere for the simulations. Under quiet geomagnetic conditions, and at middle latitudes, models can give adequate results for most applications, and are useful for computer simulation studies. However, at high latitudes, the ionosphere is highly variable and models cannot give an accurate representation of the ionosphere. Furthermore, at middle latitudes during disturbed conditions, the electron concentration makes large excursions from it's mean behaviour, and models often have to incorporate real-time data in these conditions.

3.2 TEC from GPS range measurements

The development of the TEC measurement is presented in this section. Starting from the Appleton-Hartree equation, expressions for the phase advance and signal delay are developed as differences between the true geometric path length and the measured range. This leads to a development of the dual-frequency TEC measurement. The calibration of the differential phase measurement to the differential code measurement is discussed.

3.2.1 Excess path length in GPS range measurements

In a vacuum, a radio signal propagates at the velocity of light, c. The ionosphere, however, is a weakly ionised plasma that is refractive to radio signals. Thus the velocity of a radio signal in a plasma is modified by the refractive index n, i.e. $v = c/n$. In the case

of the ionosphere, the refractive index can be determined from the Appleton-Hartree equation:

$$n^2 = 1 - \frac{X}{1 - jZ - \left[\frac{Y_T^2}{2(1 - X - jZ)}\right] \pm \left[\frac{Y_T^4}{4(1 - X - jZ)^2} + Y_L^4\right]^{\frac{1}{2}}} \tag{3.3}$$

The symbols X, Y and Z are defined as:

$$\begin{aligned} X &= \omega_p^2/\omega^2 \\ Y &= \omega_B/\omega \\ Y_L &= \omega_L/\omega \\ Y_T &= \omega_T/\omega \\ Z &= \nu/\omega \end{aligned}$$

where ω_p is the angular plasma frequency, ω_B is the electron gyrofrequency, ω_L and ω_T are the longitudinal and transverse components of ω_B respectively, and ν is the electron-neutral collision frequency. In general, n is complex, i.e. $n = \mu + j\chi$. If the refractive index n of the medium depends on the frequency of the signal passing through it, the medium is said to be dispersive.

Equation 3.3 can be simplified by making the following assumptions. First, at ionospheric heights the neutral gas density is low and there are few collisions between electrons and neutrals, that is $\nu \approx 0$. Thus absorption can be assumed to be negligible and so $Z \approx 0$. Second, the magnetic field can be assumed to be negligible, so $Y = Y_L = Y_T = 0$. Thus taking the real part of n (i.e. $\Re(n) = \mu$) Equation 3.3 reduces to

$$\mu^2 = 1 - X \tag{3.4}$$

where $X = \omega_p^2/\omega^2$. Neglecting higher order terms, which amount to less than 1% of the first-order term at GPS frequencies (see J. A. Klobuchar in Parkinson *et al.* (1996)), μ can be approximated as

$$\mu \approx 1 - \frac{1}{2}X = 1 - \frac{1}{2}\frac{\omega_p^2}{\omega^2} \tag{3.5}$$

or, since $f = 2\pi\omega$

$$\mu = 1 - \frac{1}{2}\frac{f_p^2}{f^2} \tag{3.6}$$

Here, f is the carrier frequency of the signal. f_p is the plasma frequency, given by

$$f_p^2 = \frac{e^2}{4\pi^2\varepsilon_0 m_e}N \tag{3.7}$$

where e is the charge on an electron ($= -1.602 \times 10^{-19}$ C), ε is the permittivity of free

space ($= 8.854 \times 10^{-12}$ F/m), m_e is the rest mass of an electron ($= 9.107 \times 10^{-31}$ kg) and N is the electron concentration in electrons m^{-3}. Substituting Equation 3.7 into Equation 3.6 and putting in the values for the constants gives

$$\mu = 1 - \frac{40.3}{f^2} N \tag{3.8}$$

where the constant 40.3 has units of $m^3 s^{-2}$.

Following the development in Kaplan (1996), Equation 3.8 can now be rewritten in terms of the phase refractive index, μ_p, and the group refractive index, μ_g

$$\mu_p = 1 - \frac{40.3}{f^2} N \qquad\qquad \mu_g = 1 + \frac{40.3}{f^2} N \tag{3.9}$$

and hence, using the fact that $v = c/\mu$, the phase and group velocities can be expressed as

$$v_p = \frac{c}{1 - \frac{40.3}{f^2} N} \qquad\qquad v_g = \frac{c}{1 + \frac{40.3}{f^2} N} \tag{3.10}$$

From Equations 3.10 it is clear that phase velocity, v_p, and the group velocity, v_g, differ from the free space velocity, c, by equal and opposite amounts. Hence, for the GPS signals, the PRN codes and navigation message are delayed and the carrier phase is advanced.

The path length difference due to the ionosphere can now be determined. Let the measured range between satellite s and receiver r be given by

$$S = \int_r^s \mu \, dS \tag{3.11}$$

and the true geometric range along the line of sight be given by

$$\rho = \int_r^s 1 \, dl \tag{3.12}$$

Then the path length difference is

$$\Delta S = S - \rho = \int_r^s \mu \, dS - \int_r^s 1 \, dl \tag{3.13}$$

Substituting the phase refractive index from Equations 3.9 into Equation 3.13 gives

$$\Delta S_p = \int_r^s \left(1 - \frac{40.3}{f^2} N\right) dS - \int_r^s 1 \, dl \tag{3.14}$$

where ΔS_p is the carrier phase advance, which is small in comparison with the true path length. Thus Equation 3.14 can be simplified by changing dS to dl and integrating

along the line-of-sight path. The carrier phase advance can now be expressed as

$$\Delta S_p = -\frac{40.3}{f^2}\int_r^s N dl \qquad \text{(in metres)} \tag{3.15}$$

Similarly, the group delay, ΔS_g, along the path between satellite s and receiver r can be expressed by substituting the group refractive index from Equations 3.9 into Equation 3.13 to get

$$\Delta S_g = \int_r^s \left(1 + \frac{40.3}{f^2}N\right) dS - \int_r^s 1\ dl \tag{3.16}$$

and hence

$$\Delta S_g = \frac{40.3}{f^2}\int_r^s N dl \qquad \text{(in metres)} \tag{3.17}$$

From Equation 2.1, $I = \int N dl$ is the total electron content (TEC) in electrons m^{-2}, so Equations 3.15 and 3.17 can be written:

$$\Delta S_p = -\frac{40.3}{f^2}I \qquad \text{and} \qquad \Delta S_g = \frac{40.3}{f^2}I \tag{3.18}$$

3.2.2 Dual-frequency TEC measurements

If the true geometric path length between the satellite and the receiver were known, the TEC could be calculated directly from Equations 3.18. The true path length is, of course, not known; indeed, it is the very quantity that is being sought. However, TEC can be measured directly by making differential code delay or differential carrier phase measurements on both GPS carrier frequencies. Following Mannucci *et al.* (1999), the four GPS range measurements can be expressed as

$$P_1 = p + \Delta S_{g,1} + \tau_1^r + \tau_1^s \tag{3.19}$$

$$P_2 = p + \Delta S_{g,2} + \tau_2^r + \tau_2^s \tag{3.20}$$

$$L_1 = p + \Delta S_{p,1} + N_1\lambda_1 + \epsilon_1^r + \epsilon_1^s \tag{3.21}$$

$$L_2 = p + \Delta S_{p,2} + N_2\lambda_2 + \epsilon_2^r + \epsilon_2^s \tag{3.22}$$

where P_i are the measured code ranges and L_i are the measured carrier phase ranges ($i = 1, 2$) on the L1 and L2 frequencies. p is a non-dispersive delay term which lumps together the true geometric range, clock offsets, troposphere delay and other non-dispersive hardware delays. $\Delta S_{g,i}$ and $\Delta S_{p,i}$ are the group and phase ionospheric delay terms, which are frequency dependent. λ_i are the carrier wavelengths, each of which are associated with an unknown integer cycle ambiguity N_i. The remaining terms, τ_i and ϵ_i, are the dispersive components of the satellite and receiver hardware delays. Dual-

frequency differencing of the code and phase ranges gives the ionospheric combinations P_I and L_I:

$$P_I \equiv P_2 - P_1 = 40.3\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right) I + B_r + B_s \tag{3.23}$$

$$L_I \equiv L_1 - L_2 = 40.3\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right) I + (N_1\lambda_1 - N_2\lambda_2) + B_r' + B_s' \tag{3.24}$$

The code range ionospheric combination P_I is absolute, but noisy. On the other hand, the phase combination, L_I, is smooth and can provide very accurate measurements of TEC changes, but is offset from the code range combination by a bias related to the unknown integer cycle ambiguity, $N_1\lambda_1 - N_2\lambda_2$. The situation is illustrated in Figure 3.2(a), which shows a typical phase arc between a satellite (PRN0005) and a receiver ('brus'). The red curve is the differential code TEC measurement and the blue curve is the differential carrier phase measurement. It can be seen that there is an offset of approximately 15 TEC units between the two arcs. A TEC measurement that is both smooth and absolute can be obtained by calibrating the L_I arc to the P_I arc using a least squares fit. This procedure, sometimes referred to as 'levelling', has been widely described in the literature (see, for example, Coco *et al.* (1991); Mannucci *et al.* (1998)).

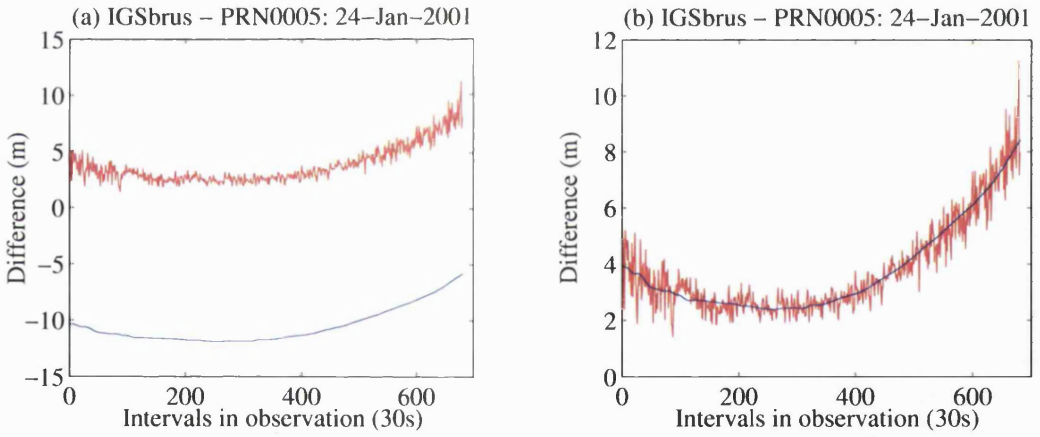


Figure 3.2: Calibration of Phase ionospheric combination (blue curve) to code ionospheric combination (red curve) for receiver 'brus' and satellite PRN0005 on 24-Jan-2001. Start time is 08:15 UT, end time is 13:55 UT. (a) before fitting phase arc to code arc, and (b) after fitting.

The resulting TEC measurement, P_I', i.e.

$$P_I' = 40.3\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right) I + B_r + B_s \tag{3.25}$$

is now both smooth and absolute, but still contains the receiver and satellite inter-frequency biases B_r and B_s, which must be estimated and removed. A description of how this is achieved is given later in this Chapter.

3.3 From TEC measurements to TEC maps

In this section, the thin shell and the tomographic approaches to mapping ionospheric TEC will be described. The two methods differ fundamentally. The thin shell model essentially models the ionosphere as a horizontal spherical shell of infinitesimal thickness, requiring vertical TECs to be estimated using a geometric mapping function. Thin shell models cannot determine the distribution of ionisation along the raypath. Tomography preserves the height dimension, allowing slant and vertical TEC measurements to be extracted directly. The tomographic method can also be extended to four dimensions by incorporating a time dependency, enabling TEC maps to be animated to show the time evolution of the TEC distribution. Whereas individual TEC measurements contain no information about the distribution of electron concentration along the line integral, ionospheric tomography offers the ability to recover this information along any ray path in the reconstruction volume.

3.3.1 The thin shell model

Maps of vertical TEC are generated from a large number of sTEC measurements between ground-based receivers and the satellites in view. The ground-based receivers are generally dual-frequency types with antennas securely fixed in precisely surveyed locations. A commonly used model, first proposed by Lanyi and Roth (1988), represents the ionosphere as a thin shell at a fixed height. The model is predicated on the assumption that the entire vertical electron density profile of the ionosphere is concentrated in a spherical shell of infinitesimal thickness. Various shell heights in the range 300 - 450 km have been reported in the literature. Referring to Figure 3.3, TEC maps are produced using the thin shell model model by the following procedure. A GPS receiver at position r makes TEC measurements to the satellites in view from its fixed position. Suppose there is a satellite at position s_1. In the first epoch, the slant TEC between r and s_1 is measured. The line of sight from r to s_1 intersects the shell (at height h) at i_1. A mapping function $M(E)$, which is a geometric function of the elevation angle E_1, is used to convert the measured slant TEC to the equivalent vertical TEC at the point p_1. Slant TEC measurements to all other satellites in the same epoch are treated similarly. All equivalent vTEC values measured in the current epoch are

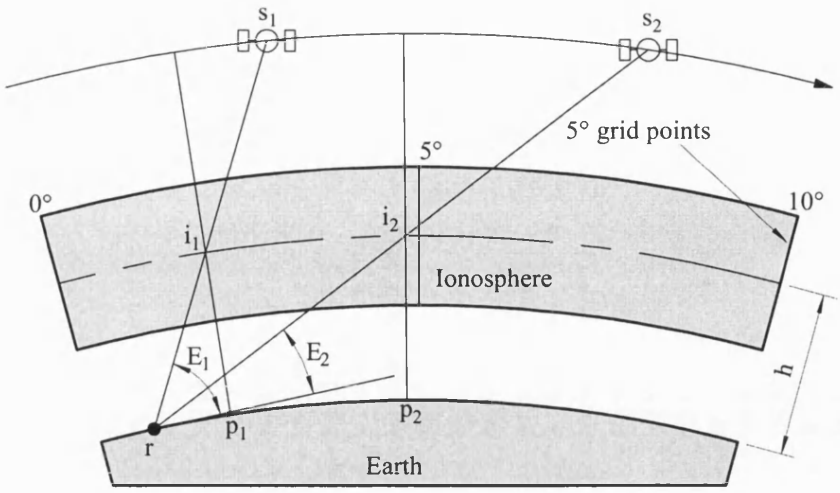


Figure 3.3: Simplified geometry of the thin shell modelling approach.

then mapped to the points of a horizontal grid using an interpolation algorithm. The grid, which is superimposed on the shell, is defined by points of latitude and longitude, typically 5° apart. Other tessellations such as equilateral triangles have been reported in the literature (Mannucci *et al.*, 1998). The original mapping function, $M(E)$, used by Lanyi and Roth (1988) is of the form:

$$M(E) = \left\{ 1 - \left[\frac{\cos E}{\left(1 + \frac{h}{R}\right)} \right]^2 \right\}^{-\frac{1}{2}} \tag{3.26}$$

where E is the angle of elevation from the receiver to the satellite, h is the shell height, and R is the radius of the Earth.

In a subsequent epoch, the satellite has moved to position s_2, and the process is repeated. Note that the elevation angle E_2 is now considerably less than E_1, so the slant path through the ionosphere from r to s_2 is much longer than the vertical path above p_2. Since TEC is a line integral, the equivalent vTEC will be overestimated at lower elevation angles. Low elevation paths are also more susceptible to multipath effects. The overestimate will be compounded by any strong horizontal TEC gradients existing above or below the shell height. An example of this effect is shown in Figure 3.4, which is a plot from an incoherent scatter radar showing electron density as a function of latitude and height. A line representing a thin shell has been superimposed on the figure at a height of 400 km. The line-of-sight from the receiver r_1 to the satellite s_1 can be seen to pass through some intense E-region ionisation at about 140 km, and some more ionisation in the F-region at 250 - 300 km. Applying the mapping function to this measurement would lead to a serious overestimate of the equivalent vTEC at the point P. Now consider the line-of-sight from the receiver r_2 to the satellite s_2, which

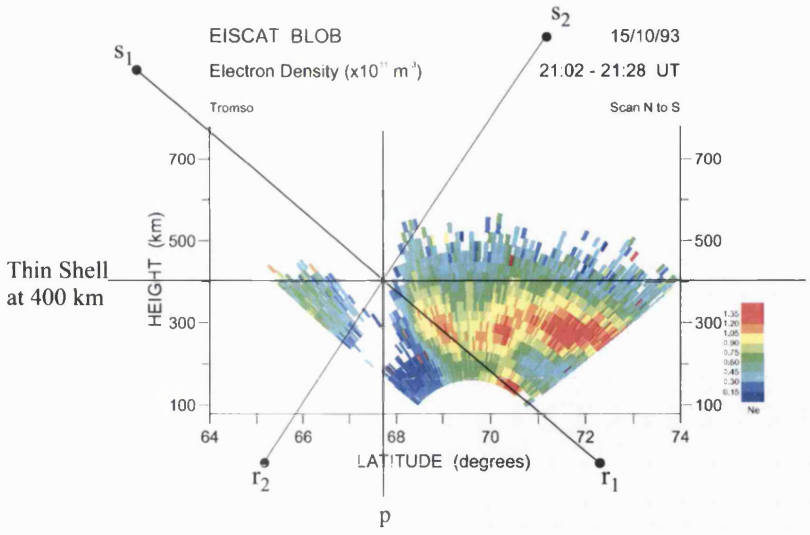


Figure 3.4: Showing the effect of strong vertical TEC gradients below the thin shell.
Source: C. N. Mitchell, PhD Thesis, 1996.

passes through some weak F-region ionisation at about 300 km, and again at 500 km. Applying the mapping function to this measurement leads to a widely different value for the equivalent vTEC at the point p. This, and other drawbacks of the thin shell model, are discussed by Davies and Hartmann (1997).

The thin shell model starts to appear in the literature around 1988. Lanyi and Roth (1988) made TEC measurements using GPS code and carrier phase data in the vicinity of a single GPS receiver, and compared them with Faraday rotation measurements from a beacon satellite by statistically fitting the TEC data to a simple model of the ionosphere based on a thin shell at 350 km. They assumed that the ionosphere is time independent for a duration of 4 - 6 hours. Using a least-squares procedure, they estimated the satellite and receiver offsets simultaneously with the TEC. Slant observations of TEC from their receiver to GPS satellites were converted geometrically to equivalent vertical TECs using the mapping function of Equation 3.26 at the points where they intersected the shell. They concluded that an upper bound for the error in GPS-measured slant TEC can be within 1 TECu compared to Faraday rotation measurements, but GPS offers superior sky coverage. However, by using a single receiver site the equivalent vTEC becomes less reliable for low elevation paths, particularly in the presence of strong horizontal TEC gradients. Also, by assuming stationarity for periods of 4 - 6 hours they would have low-pass filtered the temporal variability of the ionosphere. Nevertheless, this is an early application of the "Thin Shell" model, which has become the basis for many wide-area TEC mapping procedures.

Coco *et al.* (1991) studied the inter-frequency biases on paths between a single receiver

and the satellites in view from it. Using the least-squares estimation method of Lanyi and Roth (1988), they examined the accuracy with which the differential delay could be measured and the stability of the biases over a five-week period. They found that the satellite biases were quite stable over a two-year period. The measured inter-frequency biases for four satellites were compared with their pre-launch values, and found in two cases to agree, but to differ significantly in the other two cases. Sardón and Zarraoa (1997) examined the stability of the satellite and receiver biases over a period of 19 months from February 1995, and have attempted to quantify them separately. They considered 19 fixed GPS receivers located in various positions throughout the world, and were interested in both the day-to-day and the long-term variability. They showed that the rms long-term variability of the satellite biases over the 19-month period was less then 0.25 ns, and in some cases less than 0.1 ns. Receiver biases were found to be sensitive to almost any change in hardware, so the rms variability was computed using the longest periods without hardware changes. They found that generally the rms variability was 0.53 ns. For the day-to-day variability, the satellite biases were generally less than 0.5 ns, and receiver biases were less than 1 ns. Subsequent studies have confirmed that the biases are sufficiently stable over a period of months that they can be treated as constants (Gao and Liu, 2002). A signal generator for calibrating inter-frequency biases in TurboRogue GPS receivers is described by Duncan *et al.* (1998).

In January 1993, Wilson *et al.* (1995) of the Jet Propulsion Laboratory(JPL) applied the thin shell model of Lanyi and Roth (1988) with 30 widely distributed fixed GPS receivers and a 350 km shell to produce wide area TEC maps for the northern hemisphere. A drawback of their approach was that it required an averaging time of 6-12 hours, limiting the time resolution of their maps to 12 hours. This work was followed up later that year with a study by Mannucci *et al.* (1993), also at JPL, who describe an improved global TEC mapping method, again using multiple GPS receivers. Their approach, which also draws upon the single site method of Lanyi and Roth (1988), represented the ionosphere by a thin shell at a height of 350 km, tesselated into 642 'spherical triangles'. Using a Kalman filtering algorithm, they were able to produce vTEC maps at intervals of less than one hour. In 1998, they reported on further developments of their approach (Mannucci *et al.*, 1998), in which they included climatological models to supplement GPS data in sparse regions. Additionally, they compared their TEC maps to vTEC measurements from the TOPEX satellite, and found that the triangular gridding could reproduce latitudinal gradients at equatorial to mid latitudes reasonably well, but they suggested that more extensive comparisons in the equatorial region should be performed. In this later study, they raised the shell height to 450 km, arguing that a lower height appears to reduce accuracy. Others have studied the problem of determining an optimum shell height - see Birch *et al.* (2002), for example.

3.3.2 Ionospheric tomography

Tomography is a technique for imaging a two-dimensional object using only one-dimensional projections along a number of intersecting paths that traverse the object. Consider the grid in Figure 3.5, which contains four unknown numbers, each in a pixel of known size. The measured quantity is the sum of the path lengths through each pixel (known) times the unknown number in the pixel. For simplicity, let the pixels in Figure 3.5 have side lengths equal to 1 unit, and let the unknown numbers in all the pixels be equal to 1. Then for projection 1, the measured quantity would be $1 \times 1 + 1 \times 1 = 2$, and for projection 3 it would be $1 \times \sqrt{2} + 1 \times \sqrt{2} \approx 2.8$, etc. Hence the summation along any arbitrary path could be computed using simple geometry. The

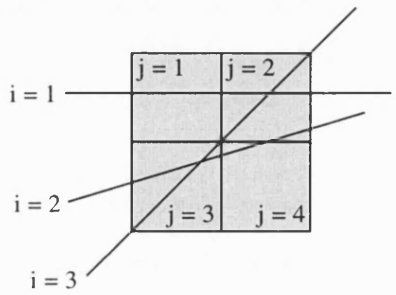


Figure 3.5: Simple grid with four pixels.

problem of tomography is to determine the unknown numbers in the pixels using only the measured projections. Let i be the indices of the projections and b_i be the $i-th$ projection, j be the indices of the pixels and x_j be the unknown number in the $j-th$ pixel. Then the length of each projection in each pixel is represented by A_{ij}, so for the grid in Figure 3.5, the following system of equations can be formed:

$$\begin{aligned} b_1 &= A_{11}x_1 + A_{12}x_2 \\ b_2 &= A_{22}x_2 + A_{23}x_3 + A_{24}x_4 \\ b_3 &= A_{32}x_2 + A_{33}x_3 \end{aligned} \tag{3.27}$$

This can be written in matrix form:

$$\begin{bmatrix} A_{11} & A_{12} & 0 & 0 \\ 0 & A_{22} & A_{23} & A_{24} \\ 0 & A_{32} & A_{33} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{3.28}$$

or, more concisely:

$$\mathbf{Ax} = \mathbf{b} \tag{3.29}$$

which could be solved for **x** using a matrix inversion, given a sufficient number of projections.

In medical tomography, the data acquisition system comprises a set of X-ray emitters aligned towards a set of sensors fixed in an assembly that is designed to rotate through 360° around the patient's body. This gives a complete set of projections **b** from all possible angles, and enables a two-dimensional image of a slice through the patient's body to be built up by solving Equation 3.29 directly. In ionospheric tomography, a grid of pixels defined by altitude versus degrees of latitude represents a vertical section through the ionosphere (see Figure 3.6). A chain of reference receivers, r_1, r_2 and r_3, placed along a meridian of longitude makes measurements of TEC along the lines of sight to a satellite s as it moves from position s_1 to position s_3. The geometry of the ionosphere and the data acquisition system severely limits the angular range of the propagation paths **b**, leading to an incomplete set of projections. Thus the matrix **A** in Equation 3.29 can be highly singular, and the system of equations cannot be solved directly by matrix inversion. *A priori* information must therefore be used to stabilise the solution before Equation 3.29 can be solved to find the underlying electron density field **x**. Note that there may be many valid solutions for **x** summing to the same line integral (i.e., the TEC).

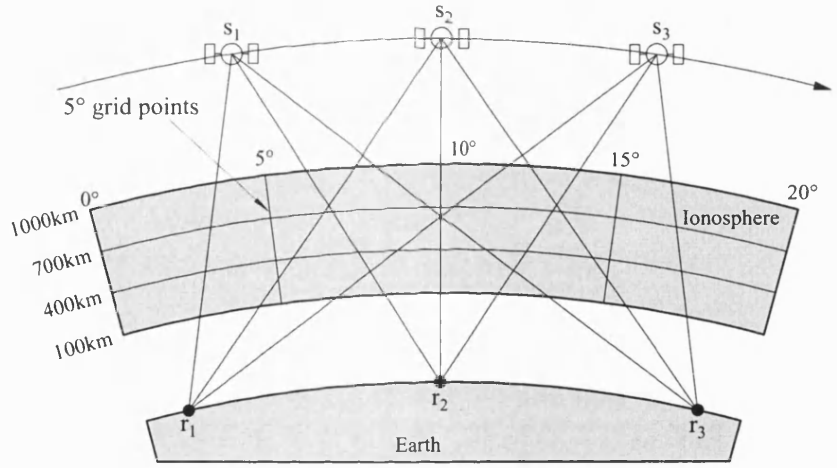


Figure 3.6: Simplified ionospheric tomography geometry.

Early experiments in ionospheric tomography used data from the NNSS satellites. Since these satellites are in polar orbits, their orbital paths would periodically be close to the longitudinal plane of the receivers. A tomographic image of a vertical slice through the ionosphere could then be made using TEC measurements from the satellites' signals on 150 and 400 MHz.

A limitation of using NNSS satellite data is that the ionosphere can only be imaged to the orbital altitude of 1100 km. However, the GPS satellites orbit at an altitude of about 20200 km, and so can provide information through almost the entire vertical extent of the ionosphere. Ciraolo and Spalla (1997) have compared TEC measurements made using GPS and NNSS data. Between January 1994 and December 1995, they made TEC observations from two NNSS stations and a GPS receiver in Italy, and found that the yearly average of GPS TEC is higher than the NNSS TEC by about 3 TEC units. They concluded that the GPS observations contain a contribution from the electron content between 1100 km and 20000 km owing to their higher orbits. In an extension of this study (Ciraolo and Spalla, 2001), they present results from 1994 to 2000, that is from low to high solar activity. Again, their results show that the average GPS TEC is higher than the NNSS TEC by 3 or 4 TEC units.

The duration of the NNSS satellites over-fly is typically about 20 minutes. The ionosphere is assumed to be stationary over that period. On the other hand, the duration of a GPS satellite can be several hours. It was initially believed that GPS would be unsuitable for ionospheric tomography, as the ionosphere could not be assumed to be stationary for the transit duration. However, since there are usually multiple satellites in view a full four-dimensional determination of the ionisation distribution can be obtained by introducing a time dependence into the inversion.

Ionospheric tomography was first proposed in 1986 by Austen *et al.* (1986). Two years later, the first tomographic images of the ionosphere, based on model data, appeared in the literature (Austen *et al.*, 1988). These were based on the case in which all raypaths lie in the same plane, i.e. the 2-D case, and imaged a region 300 km high by 2500 km wide. They used an established iterative algorithm to solve for electron concentration, producing images with six pixels vertically and twenty horizontally. By comparing images generated from their tomographic reconstructions with those from the original model ionospheres, they obtained good correlation between the two within the limitations of their geometry. Further simulation work was reported in 1990 by Raymund *et al.* (1990), in which they successfully reconstructed an ionospheric trough that had been previously mapped using incoherent scatter radar. The original image and the reconstructed plots were compared and found to agree well.

In 1991 some of the theoretical limitations to ionospheric tomography were investigated by Yeh and Raymund (1991), who show that the main limitation for ground based observations is that the geometry only permits TEC measurements to be taken over a limited range of view angles. In addition, low elevation angles are subject to multipath errors. The vertical resolution of the electron gradient is severely limited by the lack of ray paths running horizontally.

Experimental results were reported in 1992 by Pryse and Kersley (1992), who produced 2-D tomographic images from TEC measurements made by two NNSS receivers. These receivers were situated at Kiruna, Sweden and Oulu, Finland. Simultaneous measurements of electron density using the EISCAT incoherent scatter radar at Tromsø in northern Norway were taken for independent validation purposes. The results from the two methods were presented as two-dimensional images of electron density as a function of height and latitude. Whilst the two receivers were displaced from each other in longitude, and with respect to the plane of the radar scan, their tomographic reconstruction showed a large-scale latitudinal electron density gradient close to a similar structure observed by the EISCAT radar. They concluded that the agreement between the two sets of results was sufficient to justify further development of the technique. The work was extended in 1993, when Kersley *et al.* (1993) published the results of a similar study in which a chain of four NNSS receiving stations were arranged approximately on a meridian of longitude in Scandinavia. The data from these receivers were processed to yield TEC values in a vertical plane. Again, the EISCAT radar was used for independent verification of the tomographic images. The resulting images showed broad agreement with EISCAT horizontally, but they found that the lack of horizontal raypaths led to incomplete data from which to reconstruct the vertical structure. To overcome this, they created a background ionosphere using the IRI90 ionospheric model, which they note makes the results depend upon the model used.

In 1995, sensing that the proliferation of ionospheric reconstruction algorithms was causing confusion in the selection of a suitable algorithm for a particular purpose, Raymund (1995) published a comparison of different reconstruction algorithms. Three actual ionospheric electron density distributions obtained from incoherent scatter radar observations were reconstructed using a variety of iterative algorithms, and the results compared with the original distributions. Pryse *et al.* (1998) compare the performance of three different tomographic algorithms, and use independent data from the EISCAT radar for verification purposes. A review of developments in ionospheric tomography is presented by Bernhardt *et al.* (1998).

Many tomographic reconstruction algorithms start from an initial guess derived from a model ionosphere and use an iterative procedure. However, in 1992 a non-iterative parameter-fitting algorithm adapted from another geophysical situation was proposed by Fremouw *et al.* (1992). Here, the vertical structure is represented by empirical orthonormal functions (EOFs) derived from ionospheric models and the horizontal structure by a Fourier power spectrum. Their method allows *a priori* information to be fitted in a least-squares sense in order to estimate the underdetermined elements. The *a priori* information can come from model data, or independent measurements from ionosondes or radar. It is this approach that has been extended into four dimensions in the MIDAS algorithm, developed by Mitchell and Spencer (2003).

3.3.3 The MIDAS algorithm

TEC measurements taken from NNSS (TRANSIT) signal observations are suitable for 2-D reconstructions since the satellites and receivers are approximately in the same plane. The reconstructions are usually plotted as a function of latitude versus altitude. The GPS satellites and ground based receivers are not necessarily in the same plane, so 2-D reconstructions are inappropriate. To make use of the data that would otherwise be lost, a 3-D geometry can be used to recover all three spatial dimensions. Over time scales of minutes, the ionosphere can be considered to be unchanging, but the data are limited due to the long over-fly times of the GPS satellites. To overcome this data from one hour are used, and a time-dependent term is introduced into the inversion. Consequently 4-D reconstructions are possible.

A new multi-instrument inversion technique, called MIDAS, was presented by Spencer and Mitchell (2001). The aim of their method is to provide a consistent method of inverting transionospheric TEC measurements, obtained from a variety of ground and space based instrumentation, to map the spatial and temporal distribution of electron density over a wide geographical area. They demonstrate the utility of the method by mapping TEC over the entire northern hemisphere for the ionospheric storm of July 2000. Their results are confirmed by independent inversions of the mid-latitude trough position, made using NNSS data recorded in Italy. Mitchell and Spencer (2002) present a further demonstration of the potential of MIDAS in using GPS data to study the ionosphere under very disturbed geomagnetic conditions.

The MIDAS algorithm is described by Mitchell and Spencer (2003), and can be summarised as follows. Observations of slant TEC are made to GPS satellites in view from a number of ground-based fixed dual-frequency GPS receivers. A grid of three-dimensional elements (called 'voxels' to distinguish them from two-dimensional pixels) is set up in a spherical volume to represent the region of the ionosphere that it is required to image. On the assumption that the electron concentration, x, within each voxel is constant, the problem can be defined as a system of linear equations, thus

$$\mathbf{Ax} = \mathbf{b} \tag{3.30}$$

in which the $\mathbf{A}$ is a sparse matrix of the path lengths in each voxel, $\mathbf{x}$ is the vector of electron densities in each voxel, and $\mathbf{b}$ is a vector containing the observed path length integrals. Since $\mathbf{A}$ is known and $\mathbf{b}$ is the quantity measured, the problem is to solve Equation 3.30 for $\mathbf{x}$. This will permit vertical TECs to be estimated for any geographical position within the coverage area.

A difficulty arises in that many voxels in the reconstruction volume are not traversed

by any satellite to receiver ray paths, hence the path lengths in these voxels (and so the corresponding entries in the matrix **A**) are zero. The inclusion of *a priori* information will help to stabilise the solution, and the approach adopted in MIDAS uses an extension into three dimensions of the method described for two-dimensional imaging by Fremouw *et al.* (1992). Since the ionosphere cannot be considered to be stationary during the pass of a GPS satellite, a time dependence is also incorporated in the solution. Using a mapping matrix **X**, Equation 3.30 is transformed from a voxel-based representation to an ortho-normal representation of the reconstruction volume. This can be expressed as:

$$\mathbf{AXM} = \mathbf{b} \tag{3.31}$$

where the unknown term is **M**, the relative contribution of the basis functions. When phase data are used, adjacent rows of the matrices **AX** and **b** are differenced to negate the effects of the unknown carrier cycle ambiguity. **M** can now be found by inverting Equation 3.31, that is

$$\mathbf{M} = (\mathbf{AX})^{-1}\mathbf{b} \tag{3.32}$$

where AX and b are now differenced values. By applying singular value decomposition to the matrix **AX**, two orthogonal matrices **U** and **V** are returned with a diagonal matrix of singular values, **w**.

$$(\mathbf{AX})^{-1} = \mathbf{V}\left(\mathrm{diag}\left(1/\mathbf{w}\right)\right)\mathbf{U}^{\mathrm{T}} \tag{3.33}$$

Hence the solution to the inverse problem is given by

$$\mathbf{M} = \left(\mathbf{V}\left(\mathrm{diag}\left(1/\mathbf{w}\right)\mathbf{U}^{\mathrm{T}}\right)\right)\mathbf{b} \tag{3.34}$$

With **M** now known, the electron densities within each voxel can then be retrieved from Equation 3.30 and Equation 3.31:

$$\mathbf{x} = \mathbf{XM} \tag{3.35}$$

The time-dependency in the inversion is implemented in a similar manner. The following system of linear equations is defined:

$$\mathbf{Dy} = \mathbf{c} \tag{3.36}$$

in which **D** is a matrix formed from the differences in ray path geometry at consecutive time intervals, **y** is the change in electron density, which is unknown, and **c** is the change in measured TEC. The problem here is to solve Equation. 3.36 for **y**. The

equation is transformed using a mapping matrix $\mathbf{X}$, giving

$$\mathbf{DXG} = \mathbf{c} \tag{3.37}$$

where $\mathbf{G}$ is the unknown term consisting of the linear changes in coefficients of basis functions. Hence the time-dependent solution can be found by applying singular value decomposition to

$$\mathbf{G} = (\mathbf{DX})^{-1}\mathbf{c} \tag{3.38}$$

to obtain

$$\mathbf{y} = \mathbf{XG} \tag{3.39}$$

By default, MIDAS maps from the voxel-based to the ortho-normal representation using spherical harmonics in the horizontal domain, and ortho-normal basis functions in the vertical domain. These ortho-normal basis functions may be derived from models such as Chapman or Epstein. See Mitchell and Spencer (2003) for a full explanation.

3.4 The inversion procedure

A general description of how the MIDAS procedure is applied, is given in this section. The MIDAS algorithm is capable of constructing maps of ionospheric TEC using either the thin shell or the full four-dimensional inversion. An outline of the inversion parameters for both techniques is given here, and the process of computing and removing the receiver and satellite inter-frequency biases is explained.

3.4.1 Full inversion procedure

In section 3.3.2, it was shown that the full inversion problem can be expressed by the system of linear equations $\mathbf{Ax} = \mathbf{b}$ (Equation 3.29), in which $\mathbf{A}$ is the matrix of the path lengths through each voxel, $\mathbf{x}$ is the vector of unknown electron densities in each voxel, and $\mathbf{b}$ is the vector of dual-frequency TEC measurements. The problem is to solve for the underlying electron density field, $\mathbf{x}$, using the measurements of slant TEC, $\mathbf{b}$ that were calculated as described above. The vertical TEC can then be extracted by integrating through the three-dimensional electron-density volume in the vertical dimension.

For both the simulated and experimental cases, the reconstruction grid used is a low resolution 3-dimensional one covering the entire globe. Generally, the grid used in this

project is divided into latitudinal increments of 1° and longitudinal increments of 4°. In altitude, the grid extends from 80 km to 1180 km in steps of 50 km. The image regions were limited in space horizontally to 30°N - 70°N in latitude and 10°W - 40°E in longitude, as shown schematically in Figure 3.7. Where grid spacings different from these were used, this is noted in the text.

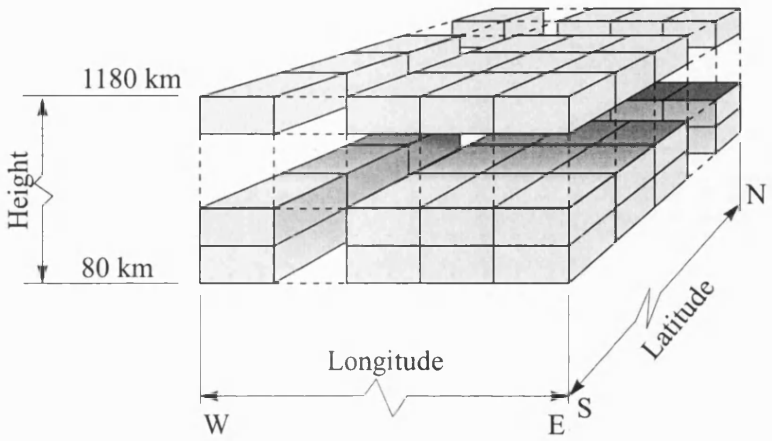


Figure 3.7: Extents of full inversion imaging grid.

The horizontal basis comprised 40 latitude harmonics and 4 longitude harmonics. In the vertical domain, three empirical orthonormal functions (EOFs) computed from eight Epstein functions are used. These have F2 electron density scale heights ranging from 30 km to 120 km, and F2 peak density heights from 250 to 400 km. The equations were solved using LU decomposition.

3.4.2 Thin shell procedure

The thin shell was approximated by a 1 metre thick layer centred in height at the shell altitude. The grid used in the thin shell images was also a low-resolution one covering the entire globe, with horizontal limits and increments equal to those of the full inversion. The imaging region was limited horizontally in the same manner as the full inversion, as shown schematically in Figure 3.8. Again, the latitudinal increments of 1° and longitudinal increments of 4° used in the full inversion were used. The horizontal constraints used were the same as those used in the full inversion, that is, 40 latitude harmonics and 4 longitude harmonics. LU decomposition was used to solve the equations. Equation 3.30 is solved for $\mathbf{x}$ as before, but it is important to note that, in this case, the matrix $\mathbf{A}$ contains no height information other than that corresponding to the thickness of the approximating layer. Thus the solution of Equation 3.30, $\mathbf{x}$, contains values of vertical TEC as a function of latitude and longitude.

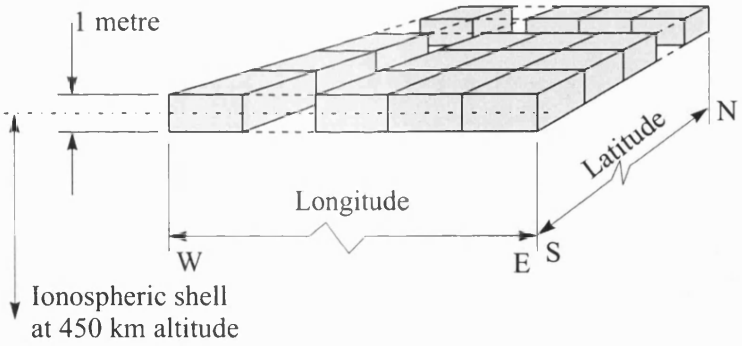


Figure 3.8: Extents of thin shell imaging grid.

3.4.3 Calculation of inter-frequency biases

In Subsection 3.2.2, the following expression (Equation 3.25) was derived for a TEC measurement that is both smooth and absolute:

$$P_I' = 40.3 \left(\frac{1}{f_2^2} - \frac{1}{f_1^2} \right) I + B_r + B_s$$

where the remaining error terms, B_r and B_s, are termed, respectively, the receiver and satellite inter-frequency biases. Due to hardware differences between the L1 and L2 RF stages, the signals propagate through the transmitters and the receivers at different, and unknown, rates. This leads to an additional error term for each satellite to receiver path equal to the sum of the inter-frequency biases. According to Sardón and Zarraoa (1997), the inter-frequency biases can amount to several nanoseconds (1 ns equates to 2.86 TECu). Typical values are given by Mannucci *et al.* (1998), who quote ±9 TECu for the satellites and ±30 TECu for certain receivers. Hence, for accurate estimates of TEC, the inter-frequency biases must be estimated and removed.

The procedure used in the MIDAS program is as follows. Each measured TEC is differenced with the corresponding TEC through the full MIDAS inversion, that is

$$\mathbf{Ax} - \mathbf{b} = \mathbf{z} \tag{3.40}$$

where $\mathbf{z}$ is the vector of residuals. Each z_i is the sum of one unique receiver bias and one unique satellite bias. Thus a system of simultaneous linear equations can be formed in which the unknowns are the individual satellite and receiver biases. So for i receiver-satellite pairs with m receivers and n satellites, the individual receiver biases,

B_r, and satellite biases, B_s, can be written

$$\begin{aligned} B_{r_1} + B_{s_1} &= z_1 \\ B_{r_1} + B_{s_2} &= z_2 \\ B_{r_2} + B_{s_2} &= z_3 \\ &\cdots\cdots\cdots\cdots \\ B_{r_m} + B_{s_n} &= z_i \end{aligned} \tag{3.41}$$

The solutions for B_r and B_s are non unique, since there are fewer independent equations than unknowns. However, each satellite-receiver bias pair is uniquely determined. More accurate values can be found by using data from longer time periods (e.g. 1 day). Thus to obtain an estimate of the true TEC, the mean daily biases must be subtracted from the TEC measurements (derived from the satellite and receiver data files). The true slant TEC can now be obtained from

$$P_I' = 40.3 \left(\frac{1}{f_2^2} - \frac{1}{f_1^2} \right) I - (B_r + B_s) \tag{3.42}$$

Figure 3.9 shows an example of the daily satellite and receiver bias calculations, computed for one of the 25 test days - 28$^{\text{th}}$ March 2001. In each subplot, the blue line represents the mean inter-frequency bias for the station computed over one day. The red points represent an estimate of the inter-frequency bias for each hour, and the green points indicate that no solution was possible. The subplot for 'bshm' (top row, 5th from left) has been reproduced in Figure 3.10 for greater detail. In Figure 3.11, the individual absolute receiver and satellite biases have been plotted for one hour, and the mean has been indicated on both plots by a magenta line parallel to the x-axis. The upper plot, which shows the receiver biases, has a mean of 19.75 TECu, and the lower plot (satellite biases) has a mean of 6.35 TECu. These values are broadly in agreement with those found in the literature (Mannucci *et al.*, 1998).

Figure 3.9: Satellite and receiver inter-frequency biases computed for 28-Mar-2001.

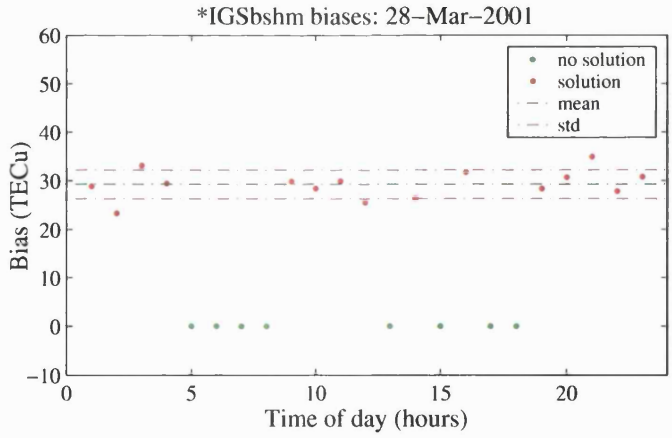


Figure 3.10: Inter-frequency biases computed for receiver 'bshm' on 28-Mar-2001.

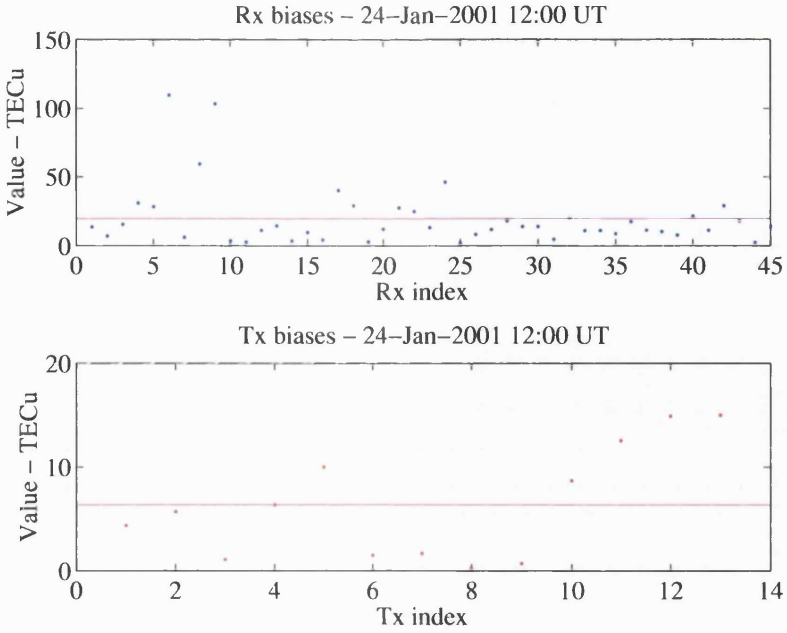


Figure 3.11: Mean inter-frequency biases computed for 28-Mar-2001 at 12:00 UT. Upper plot shows the receiver biases, with a mean of 19.75 TECu. Lower plot shows satellite biases, with a mean of 6.35 TECu.

Summary

In the first section of this Chapter, an introduction of TEC measurement and mapping was presented. The thin shell and the full four-dimensional inversion methods of TEC mapping were introduced. Some applications of TEC mapping were then discussed. In particular, the use of TEC mapping to compute corrections for the ionospheric signal delay error due for subsequent broadcast to single-frequency GPS users was discussed. This is an important application that will permit the use of single-frequency GPS receivers in civil aircraft navigation. Other applications of TEC mapping were discussed, such as those involving the design and analysis of radio communications links and remote sensing of the Earth's near-space environment for pure scientific research. The role of analytic models developed from remote sensing techniques was also discussed. In the second section, the development of the TEC measurement was considered, starting from the Appleton-Hartree equation. An expression for a TEC measurement that is both absolute and noise-free was developed, although it was shown that the inter-frequency biases remain. The two methods of combining TEC measurements to form maps of vertical TEC, the thin shell algorithm and the full four-dimensional inversion, were introduced. In the final section, a general description is given of the procedure for using the MIDAS algorithm for constructing maps of ionospheric TEC by thin shell and full inversion. The removal of the satellite and receiver inter-frequency biases is also discussed.

Chapter 4

Simulations of Thin Shell and Inversion - a preliminary study

Abstract

In this Chapter, a short study using a simulated ionosphere generated from the IRI model is presented. Using the real locations of GPS satellites and fixed dual-frequency GPS receivers, reconstructions are made for three days using the thin shell and full inversion techniques. The purpose of this is to determine whether the inversion method, which makes use of the height dimension, is capable of giving a more accurate slant TEC estimate than the thin shell method, which does not. Two European locations, Milan and Hamburg, were chosen to represent the positions of single-frequency GPS users and the slant TEC is computed between each user position and the satellites in view. Taking the model ionosphere as 'truth', the slant TECs computed using both techniques are compared in order to quantify the TEC errors. In each case, the full inversion produced smaller errors than the thin shell method.

4.1 Method

The method adopted in this Chapter is that of Meggs *et al.* (2004). Three days in September 2000, which are listed in Table 4.1, were chosen for geomagnetic conditions ranging from 'quiet' to 'disturbed', and the Kp indices obtained from the World Data Centre at www.wdc.rl.ac.uk. A model ionosphere for each day was simulated using the IRI model (Bilitza, 1990). The model inputs were the latitude, longitude and altitude

Table 4.1: Geomagnetic conditions for September 2000 days.

Date	3-hourly Kp indices (source: www.wdc.rl.ac.uk)								Ap
	03	06	09	12	15	18	21	24	
09 Sep 2000	3	2	2	2	2	2	1	2	7
24 Sep 2000	1	0	1	2	3	3	3	3	10
30 Sep 2000	3	4	6	6	5	5	5	5	45

limits, time, IG12 ionospheric index and sunspot number (RZ12), and the output of the model is the vertical electron density profile at the specified location. An example showing the model ionosphere on the Greenwich longitude for 24 September is shown in Figure 4.1. A large-scale depletion in electron density known as the main, or mid-

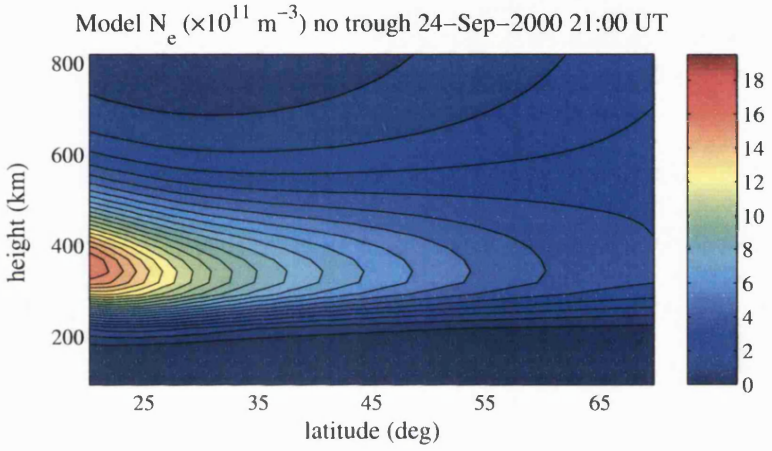


Figure 4.1: A longitudinal cut through the IRI model at 0°longitude, for 24 September at 21:00 UT.

latitude, trough is found in the night-time ionosphere at mid latitudes. Since it is not shown in the IRI model, a dynamic representation of the trough was superimposed on the IRI model ionosphere for the night-time simulations. The trough was modelled using an exponential function, shown in Equation 4.1, to scale the electron density field of the model.

$$N_{\text{sim}} = N_{\text{model}} \left[1 - d \exp\left(-\frac{(\Theta - \Theta_T)}{w} \right)^2 \right] \tag{4.1}$$

Here, N_{sim} is the electron density to be used in the simulation, N_{model} is the electron density field computed by the IRI model, d is the maximum depth of the electron density depletion inside the trough, w is the trough width, Θ is the geographical latitude and Θ_T is the trough latitude. d was set to a value of 0.5, meaning that the minimum electron density inside the trough will be half that at the same latitude in the absence of a trough. To model the steeper electron-density gradients that are associated with

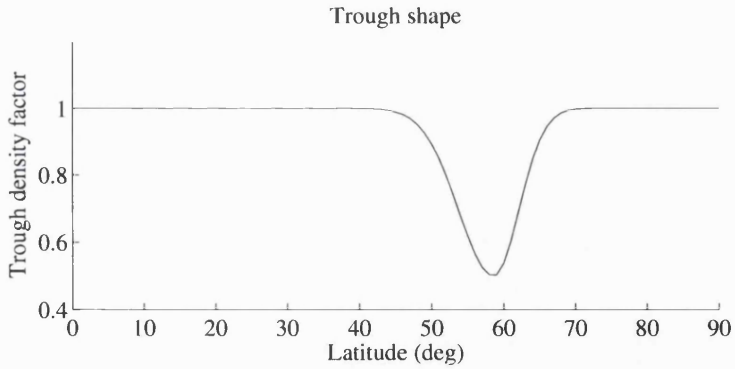


Figure 4.2: The model trough density factor plotted as a function of latitude.

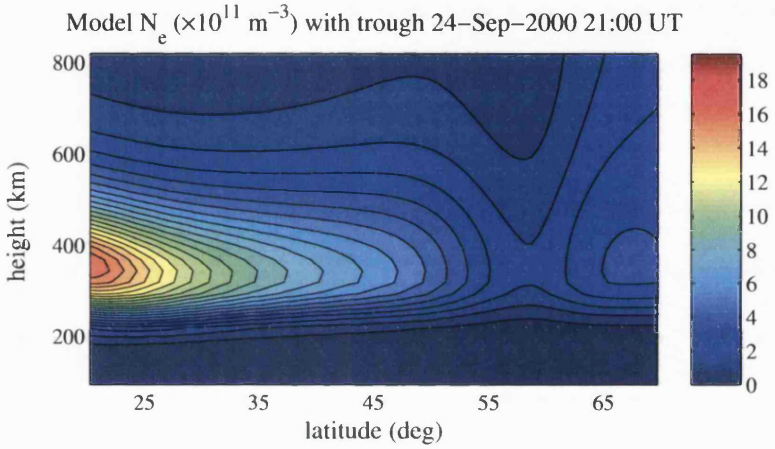


Figure 4.3: A longitudinal cut through the IRI model at 0°longitude, for 24 September at 21:00 UT. The trough is the obvious depletion of electron density at about 59°N.

the auroral region, the trough was made to be asymmetric using the following rule:

$$\text{For } \Theta < \Theta_T : \; w = 7^\circ$$
$$\text{For } \Theta > \Theta_T : \; w = 5^\circ$$

Figure 4.2 shows the resulting shape of the trough as a function of latitude. The position of the night-time trough minimum was defined using Equation 4.2, taken from Collis and Haggstrom (1988):

$$\Theta_T = 62.2 - 1.6Kp - 1.35t \tag{4.2}$$

where Θ_T is the trough latitude as before, Kp is the geomagnetic index and t is the local time difference in hours from midnight. An example of the resulting model ionosphere with the trough included is shown in Figure 4.3. Twelve GPS receivers situated in the European locations shown in Figure 4.4 were selected from the IGS (Interna-

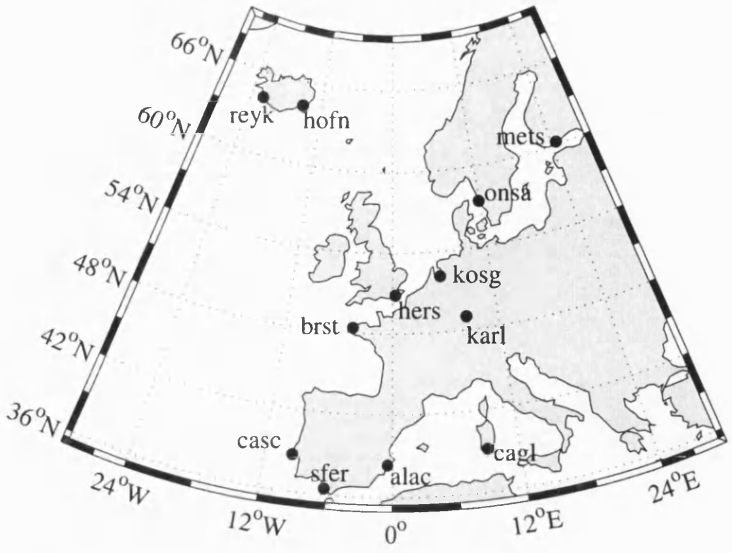


Figure 4.4: Locations of fixed GPS receivers used in the simulation study.

tional GPS Service), and the receiver observation and satellite orbital data files for all 3 days were obtained from the Scripps Orbit and Permanent Array Center (SOPAC) archive at http://sopac.ucsd.edu/. All GPS satellites in view from the receivers were used for the simulation, subject to a minimum elevation limit of 15° from each receiver location. No account was taken of the local topography at each site. TEC observations were simulated by integration through the IRI model between each GPS satellite and receiver, and the resulting model ionospheres were taken as the truth to which the simulated reconstructions were related. The inter-frequency biases, multipath and tropospheric sources of error were neglected. The MIDAS program was then used to generate vertical TEC maps from the simulated observations at a range of horizontal resolutions, for both thin shell and full inversion. The inversion is an under-determined problem due to the sparse nature and geometrical bias of the GPS satellite-to-ground measurements, and so the solution must be stabilised. The approach adopted here was to use a model in which the vertical structure is represented by an orthogonal basis set derived from a range of modelled conditions or mathematical functions, and the horizontal structure is represented by spherical harmonics. Chapman functions were used in this study, although the basis set can be formed from any ionospheric model. See Chapter 3, Subsection 3.3.3 for a description of the MIDAS algorithm.

The full inversions encompassed a height range of 80 km to 1180 km in increments of 50 km, and the thin shell reconstructions used a shell height of 400 km. Identical time evolution terms were implemented in both methods as described in Mitchell and Spencer (2003). Both techniques were applied using five different horizontal resolutions (i.e. grid spacings) ranging from 1° to 5° in steps of 1°. Two test locations, Milan at

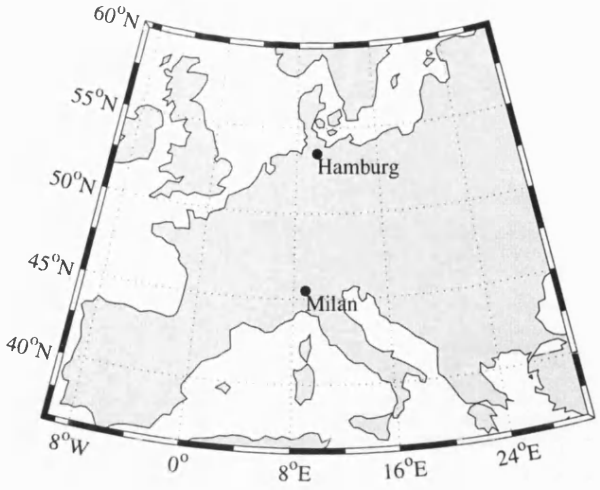


Figure 4.5: Locations of user receivers at Milan and Hamburg.

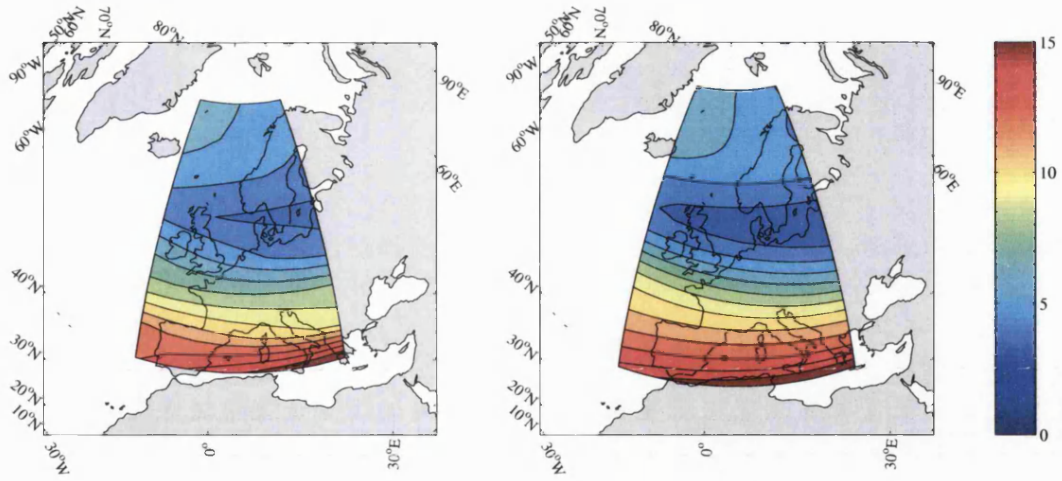
45.5° N, 9° E and Hamburg at 53.5° N, 10° E (see Figure 4.5), were chosen to represent single-frequency GPS users who require real-time ionospheric corrections. The accuracy of the TEC determination in the two techniques was assessed by comparing the true slant TEC (obtained by integration through the original model) from the test locations with the corresponding slant TEC using Equation 4.3:

$$\text{Error} = \sqrt{\frac{\sum_{i=1}^{N_p} (I_m - I_c)^2}{N_p}} \tag{4.3}$$

where I_m is the simulated truth TEC, I_c is the TEC computed from either the thin shell or the inversion method, and N_p is the number of satellite to ground paths.

4.2 Results

Two maps of vertical TEC for the late evening (21:00 to 22:00 UT) of the disturbed day, 30th September 2000, are shown in Figure 4.6. The model ionosphere gave a mean slant TEC of 19.1 TEC units (TECu - 1 TECu is equivalent to 10^{16} electrons m^{-2}) over all ray paths from the Hamburg test location in this period. In Figure 4.6(a), the TEC map has been generated using a thin shell at 400 km, with a grid size of 5° in both latitude and longitude. The mean slant TEC in the thin shell was 18.7 TECu, so the thin shell has under-estimated the TEC. Figure 4.6(b) shows a TEC map generated from a full inversion, with a grid size of 1° in latitude and longitude. In this case, the mean slant TEC was 19.3 TECu, i.e. an over-estimate of TEC. In each case that was



(a) (b)

Figure 4.6: Two maps of vertical TEC for a user at Hamburg on 30 September at 21:00 UT., generated from (a) thin shell at 400 km, and (b) full inversion. The colour bar is common to both maps.

examined, the absolute value of the inversion TEC error was found to be lower than the thin shell TEC error.

In Figure 4.7, separate graphs are shown for TEC errors from Milan and Hamburg, and for daytime and night-time. The horizontal grid spacing in degrees is shown on the x-axis, and the y-axis shows the error in metres computed by Equation 4.3. The thin shell errors are represented by diamonds and the full inversion by squares. The results show that, as the horizontal resolution is decreased, the accuracy is improved, as expected. They also indicate that the full inversion method is capable of giving an overall improvement in accuracy compared to the thin shell.

A further exercise concerns the optimal height for the shell, about which there is much debate in the literature. Birch *et al.* (2002) present a discussion in which they suggest that the optimum height should be as much as 750 km - 1200 km in order to allow for the protonospheric contribution. More commonly used figures are 350 km (Mannucci *et al.*, 1993), 400 km (Ciraolo and Spalla, 1997) and 450 km (Mannucci *et al.*, 1998). The procedure used above was carried out for a range of shell heights and the slant TEC error was plotted in Figure 4.8. It can be seen that, in this case, the error is a minimum when the shell height is set to 400 km.

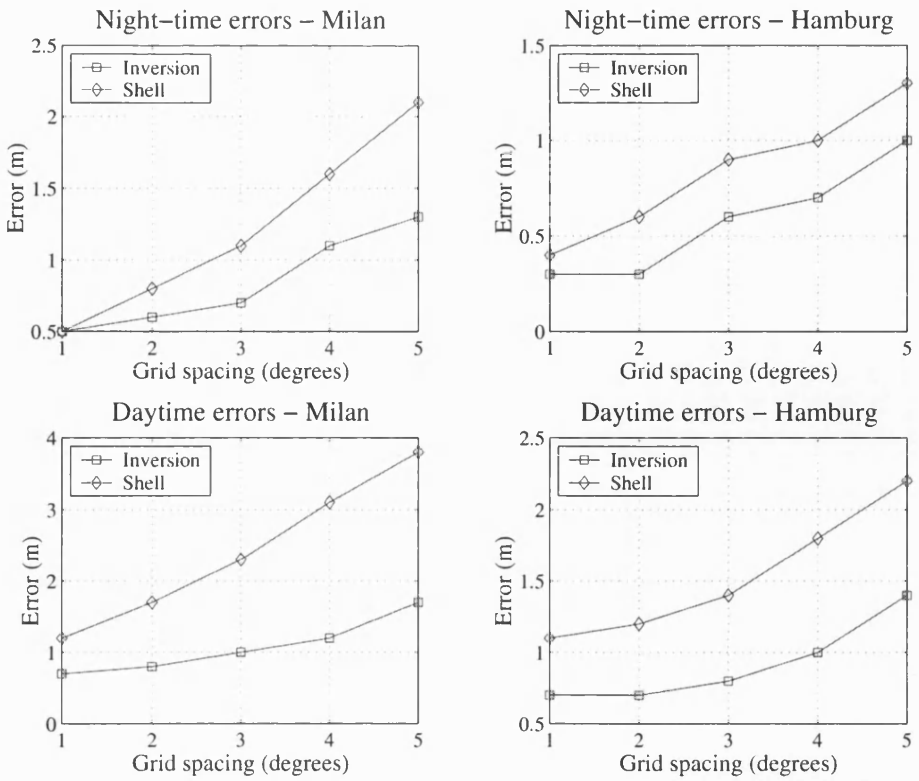


Figure 4.7: The errors in metres for the thin shell and the full inversion. The upper two plots show the night-time errors, and the lower two show daytime errors.

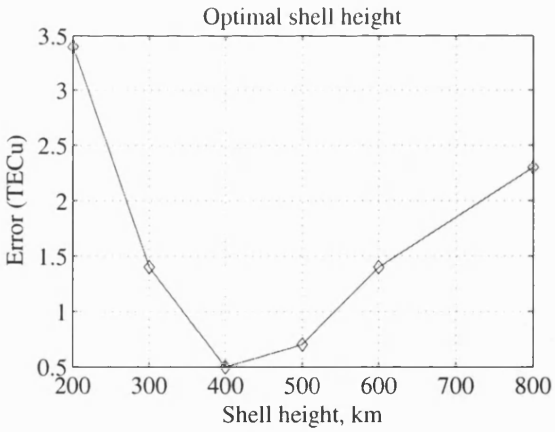


Figure 4.8: TEC errors for a range of thin shell heights.

4.3 Discussion

The simulation work presented here has some shortcomings that may offer advantages for both TEC mapping techniques. The simulated ionosphere used here is based on an empirical model, which contains no small-scale features of less than 100 km. The IRI grid size is 4° in longitude, 1° in latitude and 25 km in height, and so there are no small-scale irregularities that can cause phase jumps in the simulated TEC records. Furthermore, multipath effects were not modelled in the simulation, and the local topography for the receiver sites was ignored. It was therefore assumed that all satellites were visible from all sites down to the 15° elevation angle.

It should be noted that, in addition to the multipath effects, the inter-frequency biases and the tropospheric delay were not taken into account. Also, straight-line propagation was assumed, since the voxels are large in comparison with the scale of the ray bending. Since all of these factors are common to the thin shell and the full inversion, their omission would not be expected to favour either technique.

The results suggest that there is an advantage in using smaller grid spacings in the reconstruction. There are two reasons for this. Firstly, the smaller grid spacing allows the voxels to be smaller, and hence the TEC integration is smoother. Secondly, reconstructions of electron density are better defined on a smaller grid. This is not simply a matter of interpolation, but because there is more information contained in the smaller grid. However, there would be no point in making the reconstruction grid smaller than the IRI grid size as there is no further information to be gained from doing so. The size of the IRI grid size therefore represents the lower limit to the size of the reconstruction grid for the simulations.

A valid conclusion of this simulation study is that the full inversion can give a better estimate of the slant TEC than the thin shell.

4.4 Summary

The feasibility of using a four-dimensional inversion technique to estimate the slant TEC has been demonstrated. Using a minimal network of European fixed GPS receivers and real satellite positions, reconstructions using the thin shell and the full inversion methods were made, and each was compared to a simulated ionosphere. The results were computed across 24 hourly inversions and five different horizontal resolutions.

This study has indicated that the four-dimensional (full) inversion of GPS data is a feasible approach to the problem of imaging the ionosphere at mid-latitudes, and offers certain advantages over the thin shell method by overcoming problems associated with mapping functions, as described in Chapter 2. In the next Chapter, a long term study is presented of the inversion method, using real GPS observation data from a more extensive network of fixed GPS receivers, and covers selected days from the whole of 2001.

Chapter 5

Simulations and experimental work - a long term study

Abstract

In Chapter 4, a short simulation study was presented in which thin-shell and full-inversion reconstructions were compared with a simulated ionosphere. An extended study is presented in this Chapter. Using observation data from an extensive network of fixed GPS receivers situated in Europe, images of the ionosphere are produced using both the thin shell and the MIDAS full inversion methods. Twenty-five days in 2001 have been selected for the work of this Chapter, which divides into three well-defined studies. In the first, a simulated ionosphere is used as truth, against which the vertical TEC obtained by the thin shell and the full inversion techniques can be compared. In the second study, thin shell and full inversion reconstructions are made using the same satellite and receiver geometry as in the first study, but with measurements of real slant TEC. A qualitative comparison is made between each technique and the input measurements. In the third study, errors in the slant TEC derived from the MIDAS images are estimated by comparing slant TEC measurements with the corresponding integrated paths through the full inversion grid of electron density.

5.1 Introduction and method

In this Chapter, an extended study is presented, in which thin shell and full inversion images of the ionosphere are made for 25 days throughout 2001, in both simulated and

experimental cases. At least one geomagnetically quiet day, and one geomagnetically disturbed day, were selected from each month by consulting the Kp indices for 2001 at the World Data Centre (www.wdc.rl.ac.uk). The selected days are listed in Table 5.1, together with with their 3-hourly Kp indices and Ap values.

Table 5.1: Geomagnetic conditions for the selected days of 2001.

Date	3-hourly Kp indices (source: www.wdc.rl.ac.uk)								Ap
	03	06	09	12	15	18	21	24	
01 Jan 2001	0	0	1	1	1	0	1	1	3
24 Jan 2001	2	3	1	4	5	4	4	3	18
04 Feb 2001	0	0	0	0	1	1	1	0	2
13 Feb 2001	2	4	4	3	3	4	4	4	19
16 Mar 2001	0	0	0	0	1	2	2	1	2
28 Mar 2001	2	2	4	6	6	4	3	4	31
11 Apr 2001	3	3	3	2	4	8	6	7	60
30 Apr 2001	1	0	0	0	1	3	3	1	5
06 May 2001	1	0	2	2	3	3	2	2	7
12 May 2001	4	2	4	6	6	4	3	4	34
31 May 2001	1	0	0	1	3	3	2	2	7
18 Jun 2001	2	5	5	6	5	3	4	4	34
28 Jun 2001	0	0	1	2	2	3	2	2	5
04 Jul 2001	2	2	1	2	3	3	2	3	8
25 Jul 2001	3	4	5	4	3	4	4	3	22
17 Aug 2001	1	1	2	4	5	5	6	5	29
24 Aug 2001	1	1	1	1	2	2	2	1	4
07 Sep 2001	1	0	1	1	3	3	1	3	6
23 Sep 2001	0	3	5	4	5	5	5	4	27
07 Oct 2001	0	0	0	1	2	2	2	2	4
22 Oct 2001	6	5	4	6	5	7	6	5	66
06 Nov 2001	8	8	7	5	6	7	5	5	112
27 Nov 2001	1	0	0	0	0	2	1	1	2
09 Dec 2001	2	0	0	1	1	1	2	1	3
24 Dec 2001	3	3	5	4	4	4	3	2	19

In the work of Chapter 4, a minimal network of twelve fixed dual-frequency GPS receivers in European locations was selected. In this study, the more extensive receiver network shown in Figure 5.1 is used. There are currently many hundreds of suitable dual-frequency receivers distributed throughout the world, and their observation files are readily available free of charge from the SOPAC archive at http://sopac.ucsd.edu. Satellite orbital data files are also available from the same source. Care was taken in the selection of the receivers to ensure a good distribution over mainland Europe

The receiver observation files contain both pseudo-code and carrier phase path length measurements, which are recorded at 30 s intervals. In the present study, the measure-

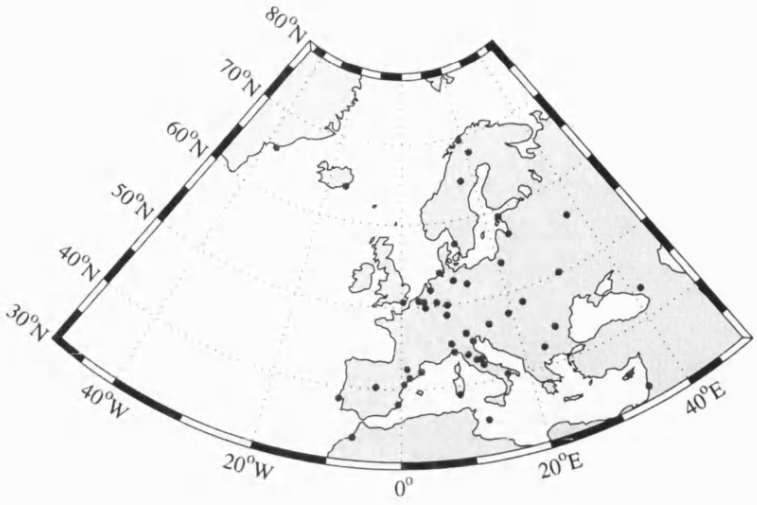


Figure 5.1: Locations of fixed GPS receivers (note that the receiver idents have been omitted for clarity).

ments were under-sampled by a factor of 4, giving one observation every 2 minutes, which is adequate for the inversion technique. The measured TEC vector, **b** (in Equation 3.30), used in the inversions was formed from dual-frequency code and carrier-phase measurements in the manner described in subsection 3.2.2. After calibrating the pseudo-code to the carrier phase code measurements, the TEC measurement is absolute and free of multipath and noise effects, but still contains the satellite and receiver inter-frequency biases.

For the simulation studies, an artificial ionosphere was constructed for each of the 25 test days, using the IRI-95 model (Bilitza, 1990). The inputs to the model are: time, geographical position, IG12 index and RZ12, and the model returns the electron density profile for the specified geographical position and time.

Using the procedure outlined in Chapter 3, thin shell and full inversion reconstructions were generated for all 25 days, in both simulated and experimental cases. The same satellite and receiver geometry is used in each case. After running the full inversion, the daily mean inter-frequency biases are calculated as described in subsection 3.4.3. These were then used to correct the TEC measurements that are used in the slant TEC study using real data.

Shell altitudes found in the literature generally range from 300 km to over 700 km, with most in the 350 - 450 km range. The choice of shell altitude used here was based on the work of the previous Chapter and on Mannucci *et al.* (1998). In Figure 5.2, Figure 4.8 from Chapter 4 is reproduced, and a cubic fit (blue curve) has been superimposed

on the original data (red data points). The cubic fit has a minimum at about 426 km. This value was rounded up to 450 km in agreement with Mannucci *et al.* (1998) for use in this study.

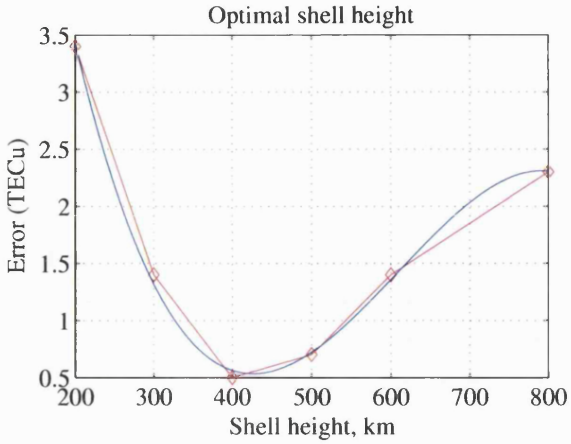


Figure 5.2: TEC errors for a range of thin shell heights, with cubic fit applied.

A vertical TEC map was plotted for each hour and each day, for the model and for the reconstructions. Then, taking the IRI-95 model ionosphere as truth, the question under consideration here is: If the true ionosphere is the one generated by the IRI-95 model, how well can that ionosphere be reproduced by (a) MIDAS inversion and (b) a thin shell model?

5.2 Vertical TEC from simulations

In this section, thin shell and full inversion TEC maps are compared with a simulated ionosphere ('truth'), which was generated using the IRI-95 model. The example results presented are for the 28th March 2001, a geomagnetically disturbed day with Kp values ranging from 2 to 6 (see Table 5.1). The inversions are carried out at hourly intervals throughout the day, and four hours have been chosen for discussion in this section. Mean vertical TEC errors for each hour in the 25-day data set are also presented.

The four hours that have been chosen for examination are 03:00 UT, 06:00 UT, 14:00 UT and 18:00 UT. The diurnal pattern of ion production and loss has been described in Chapter 2, but a brief recap is given here. During the night-time, the absence of solar radiation means that there are no electron production processes active, but recombination and other electron loss processes continue regardless. Electron concentration falls to a minimum at about 03:00 local time, at which time average vertical TEC values

Table 5.2: Local time at longitudinal edges and centre of images.

UT	Local time at longitude		
	10°W	15°E	40°E
03:00	02:20	04:00	05:40
06:00	05:20	07:00	08:40
14:00	13:20	15:00	16:40
18:00	17:20	19:00	20:40

can be expected to be around 3 TECu. At sunrise, solar radiation starts to trigger production of electrons, and strong horizontal TEC gradients can be expected as the ionosphere transits from the night-time regime to the daytime. Since 28th March is close to the vernal equinox, 06:00 UT was taken as an appropriate time for the sunrise period. Ionisation continues to increase during the morning and early post-noon periods, causing TEC levels to reach their maximum value between 14:00 and 15:00 local time. The intensity of the ionising solar radiation decreases during the afternoon and the transition from the daytime regime to the night-time commences at sunset. During this period, levels of TEC start to decay and strong horizontal TEC gradients can be expected. Again, with 28th March being close to the vernal equinox, 18:00 UT has been chosen. The TEC maps are all identified by universal time (UT). The equivalent local time for the western, central and eastern longitudes of the image region are given in Table 5.2.

5.2.1 Vertical TEC - 28 March 2001 at 03:00 UT

The first example results can be seen in Figure 5.3, which shows four TEC maps for 28th March 2001 at 03:00 UT. The Kp index at this time was about 2, indicating low geomagnetic activity. The IRI-95 model ionosphere (truth) is shown in Figure 5.3(a), which exhibits vertical TEC values ranging from about 12 TECu over the southern Baltic Sea to about 30 TECu in the extreme south-west of the imaged region. Figure 5.3 (b) shows the equivalent vertical TEC at the points at which the receiver-to-satellite raypaths intersect with a 450 km shell. Since the plot is generated from the measured slant TEC, corrected for inter-frequency biases as described in Chapter 3, it represents the average state of the ionosphere over a 1-hour period, and hence is a useful visual check that the inversion is a reasonable representation of the ionosphere. The sampling interval is 2 minutes, thus there is a maximum of 30 points per hour for any receiver to satellite path.

The TEC map at Figure 5.3 (c) has been produced using the full MIDAS inversion. A

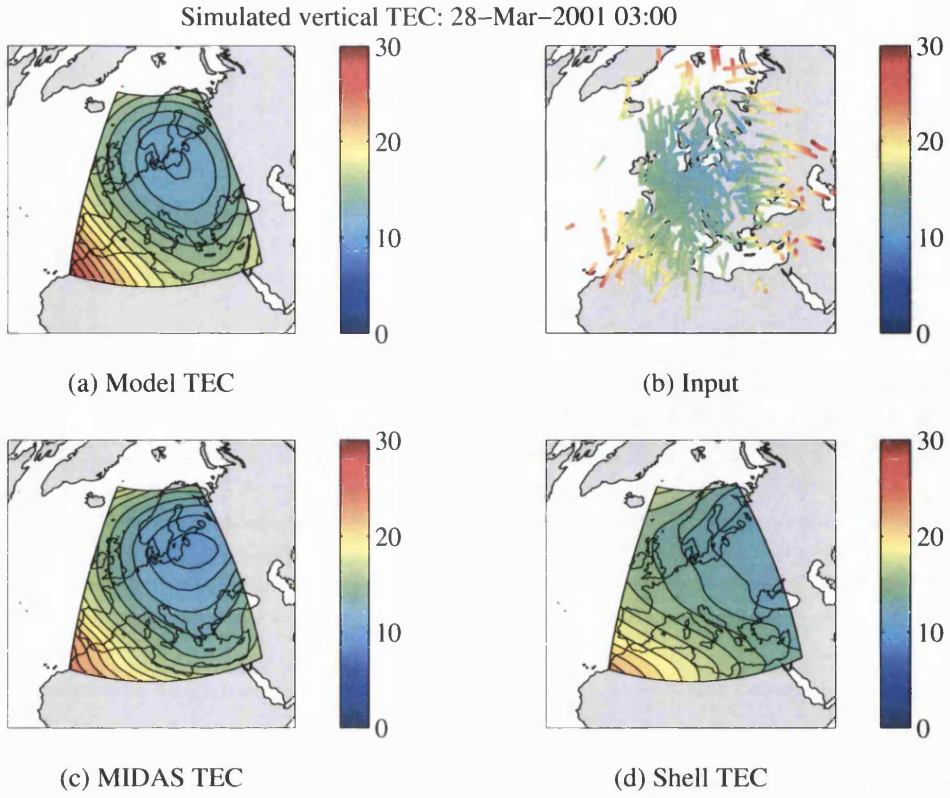
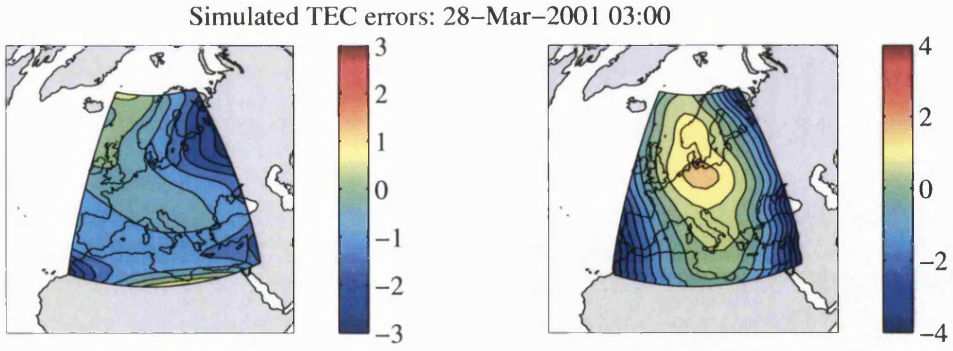


Figure 5.3: TEC maps for 28 March 2001 at 03:00 UT: (a) Model ionosphere from IRI-95, (b) input to reconstructions, (c) vertical TEC map from MIDAS and (d) vertical TEC map from thin shell.

visual inspection reveals that there is good overall agreement with the model ionosphere at Figure 5.3 (a). Closer inspection reveals that the inversion has slightly underestimated the vertical TEC value near the minimum over southern Sweden and Poland. The values of vertical TEC can be seen to range from about 10 TECu over the southern Baltic Sea, to a maximum of about 25 TECu in the south western corner, with some distortion of the contours evident near the TEC maximum. A comparison with the input plot at (b) further confirms that the TEC measurements through the model have produced a reasonable representation of the ionosphere.

Figure 5.3 (d) shows a vertical TEC map of the model ionosphere produced using a thin shell at 450 km. It is immediately clear from visual comparison with the model that the thin shell has not reproduced the ionosphere as closely as did the full inversion. In particular, the steep TEC gradients in the south western corner of the model have not been well reproduced by the thin shell model. The TEC minimum over the southern Baltic is also less well defined. TEC values can be seen to range from about 12 TECu in the east to about 21 TECu in the south west.



(a) MIDAS Errors

(b) Shell Errors

Figure 5.4: Contour plots of absolute errors in TECu between (a) MIDAS inversion and the model, and (b) thin shell and the model.

In order to quantify the errors in the reconstructions, the vertical TEC values computed by the MIDAS inversion and the thin shell were each differenced with the values of vertical TEC in the original model. Figure 5.4 shows the distribution of the absolute errors over the imaged region. These were computed from $error = TEC_{image} - TEC_{model}$. Green contours indicate small errors, and hence the reconstruction has computed the model vertical TEC correctly. Positive error values (in the red direction) indicate that the reconstruction has overestimated the model vTEC, and negative values (blue direction) indicate that the reconstruction has underestimated the model vTEC. Note that the colour range for the MIDAS errors is from -3 to +3 TECu , and for the thin shell errors the range is from -4 to +4 TECu .

In Figure 5.4 (a), the errors in the MIDAS inversion for 28^{th} March 2001 at 03:00 UT can be seen. In general, the MIDAS inversion has underestimated the model vTEC by about 1 to 2 TECu . There are small regions in which the model vTEC has been overestimated by less than 1 TECu unit, but these are very close to the edges of the imaged regions and may have been caused by inadequate receiver coverage at the extremes of the imaged region. Comparing this with the thin shell errors shown in Figure 5.4 (b), it can be seen that the errors are generally slightly larger, ranging from an underestimate of about 4 TECu to and overestimate of about 2 TECu . It is also noticeable that the thin shell TEC error map exhibits more variability than the MIDAS inversion TEC error map.

5.2.2 Vertical TEC - 28 March 2001 at 06:00 UT

The TEC maps for 28^{th} March 2001 at 06:00 UT are shown in Figure 5.5. As this is the sunrise period at this time of the year, sharp horizontal TEC gradients can be expected as the night-time ionosphere transits to the day-time. The geomagnetic activity is still low, as indicated by the Kp index of 2 (Table 5.1).

Following the same pattern as in subsection 5.2.1, the simulated TEC (i.e., 'truth') is in Figure 5.5 (a), and the input plot is in Figure 5.5 (b). The simulated ionosphere displays a smooth west-east gradient, with TEC ranging from about 10 TECu in the west to about 40 TECu in the south-east. This is in agreement with the input plot. The sun would be at about 90°E at this time, so the elevated TEC levels in the south-east are as expected.

The inversion TEC maps are in Figure 5.5 (c) and (d). Comparing these qualitatively with the simulated ionosphere, it is immediately apparent that the full MIDAS inversion has reproduced the TEC distribution much more closely than the thin shell. Closer inspection of the full inversion TEC map reveals some distortion of the contour lines, particularly along the southern boundary of the image region. This may be due to unfavourable satellite-receiver geometry. However, this distortion is minor in comparison with the thin shell TEC map, in which the smooth west-east gradient observed in the model has not been well reproduced. Indeed, the thin shell has seriously overestimated the TEC in the west of the imaging region, underestimated the TEC in the east, and seriously distorted the TEC contours.

The absolute errors in vertical TEC, shown in Figure 5.6, were computed in the same way as those in the previous subsection. It can be seen from Figure 5.6 (a) that the full MIDAS inversion has produced absolute TEC errors that are generally less than 30% of those in the thin shell. Generally, the full inversion method has underestimated vertical TEC over mainland Europe by less than 2.5 TECu. Over Scandinavia, vertical TEC has been overestimated by 1.5 TECu or less. On the other hand, Figure 5.6 (b) reveals that the thin shell has overestimated the vertical TEC in the west of the imaging region by about 10 TECu, and in the east it has underestimated vertical TEC by about 5 TECu.

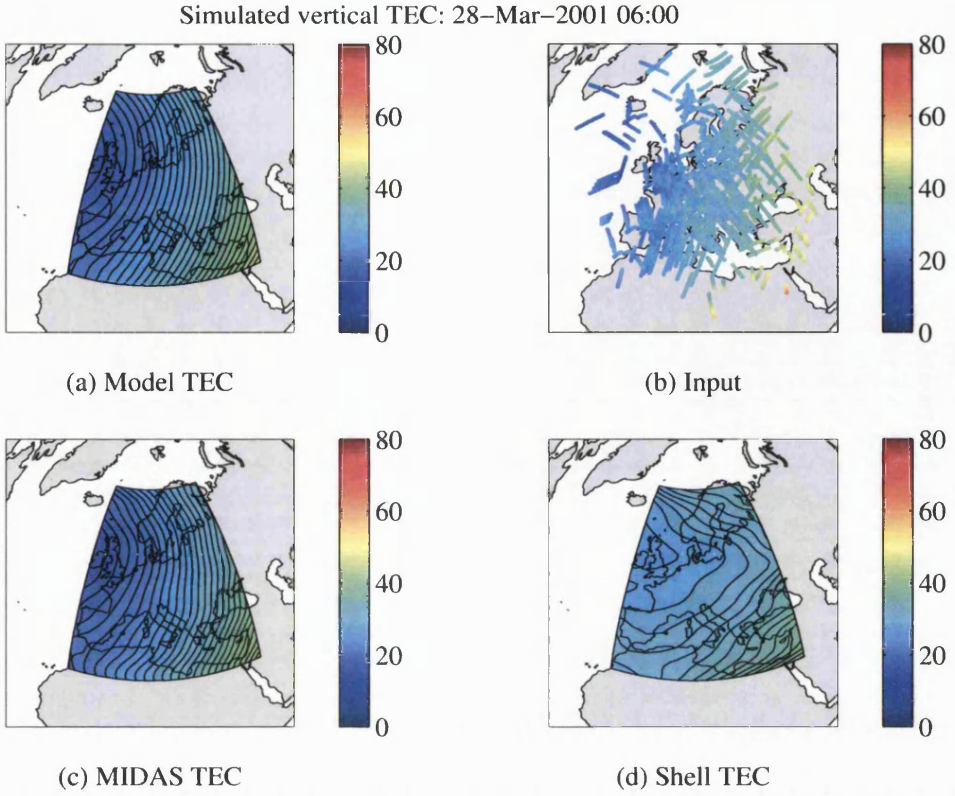


Figure 5.5: TEC maps for 28 March 2001 at 06:00 UT: (a) Model ionosphere from IRI-95, (b) input to reconstructions, (c) vertical TEC map from MIDAS and (d) vertical TEC map from thin shell.

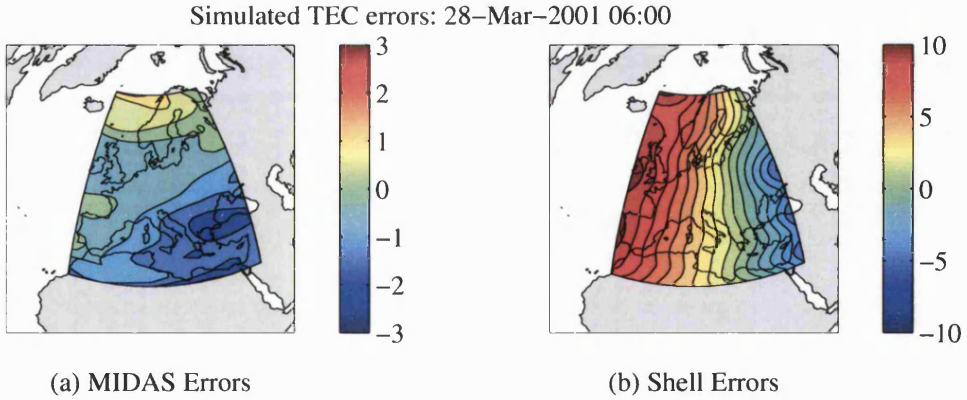


Figure 5.6: Contour plots of absolute errors in TECu between (a) MIDAS inversion and the model, and (b) thin shell and the model.

5.2.3 Vertical TEC - 28 March 2001 at 14:00 UT

The next example (see Figure 5.7) shows the four TEC maps for 28th March 2001 at 14:00 UT. This is the time of day at which TEC levels are likely to be at their maximum. Furthermore, the Kp index at this time was about 6, indicating moderately high geomagnetic activity. The simulated ionosphere generated from the IRI-95 model is shown in Figure 5.7 (a), and shows TEC levels to be elevated compared to those at 03:00 UT (see Figure 5.3) due to the diurnal variation in TEC. A steep TEC gradient can be seen over North Africa and the Mediterranean, which is reproduced quite faithfully in the MIDAS TEC map of Figure 5.7 (c). The region of highest TEC is to the south west of the image, which is consistent with the expected position of the sun at 14:00 UT. There is also good agreement between the model and the MIDAS image with respect to the structure in the east of the imaging region. This structure is evident in the input plot of Figure 5.7 (b). To the north and west of the MIDAS image some artificial structure is evident, which may be due to poor receiver coverage in this area.

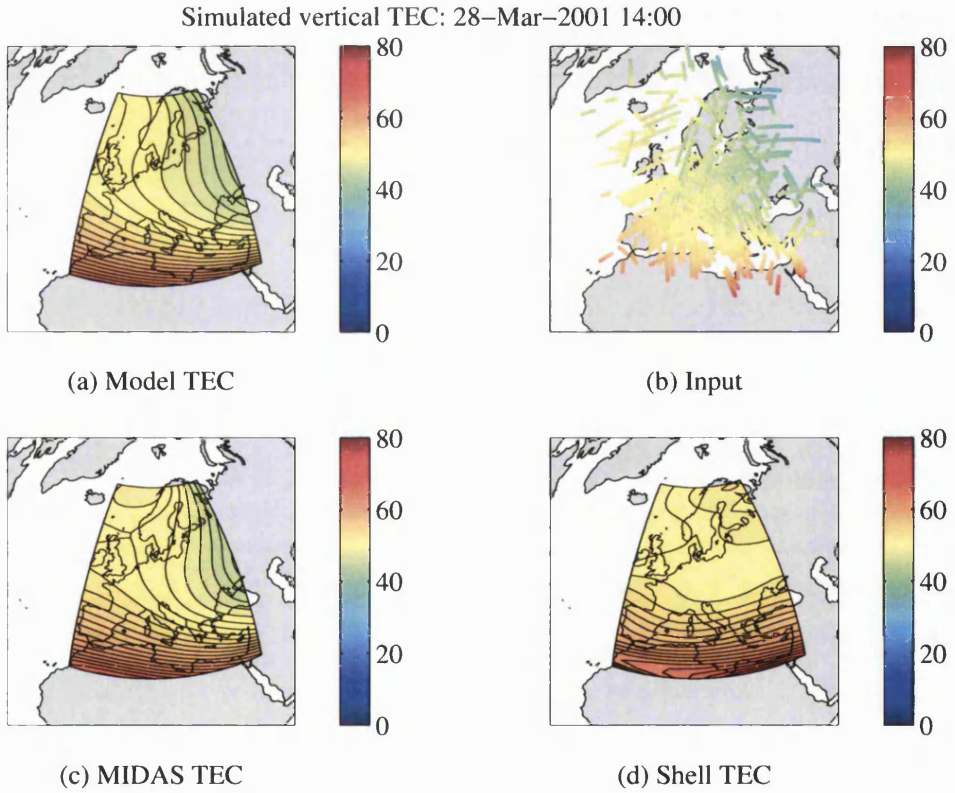
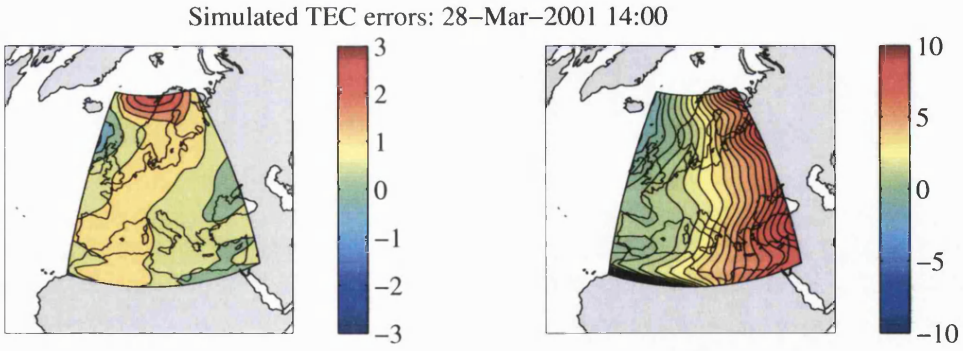


Figure 5.7: TEC maps for 28 March 2001 at 14:00 UT: (a) Model ionosphere from IRI-95, (b) input to reconstructions, (c) vertical TEC map from MIDAS and (d) vertical TEC map from thin shell.



(a) MIDAS Errors

(b) Shell Errors

Figure 5.8: Contour plots of absolute errors in TECu between (a) MIDAS inversion and the model, and (b) thin shell and the model.

The thin shell TEC map is shown in Figure 5.7(d). It is immediately apparent that the vertical TEC distribution is less well represented than in the MIDAS inversion image. The steep TEC gradient seen in the south of the IRI-95 model is also evident in the thin shell reconstruction, but the thin shell has not reproduced the gradient as closely as the MIDAS image. Some levelling out of the gradient can be seen at the southern edge. Also, the structure to the east of the region is entirely absent in the thin shell image. Some distortion of the contours is also evident in the over the United Kingdom and the North Sea.

The vertical TEC errors in the MIDAS and the thin shell images are plotted as contours in Figure 5.8. Here again, the errors were computed using $error = TEC_{image} - TEC_{model}$. In both plots, green regions indicate small error, red regions indicate that the model has been overestimated by the image, and blue regions indicate an underestimate of the model. In this case, the colour range for the MIDAS errors is from -3 to +3 TECu , and for the thin shell errors the range is from -10 to +10 TECu .

Taking the MIDAS inversion errors first (see Figure 5.8(a)), the artificial structure that was evident in the MIDAS TEC map above appears as a region of large positive error - in the order of 3 TECu. This would be equivalent to a path length error of about 0.5 m on the GPS L1 frequency. Elsewhere, the MIDAS inversion image has generally overestimated the model by about 1 TECu unit (0.163 m at L1). A small underestimate of less than 1 TECu unit can be seen between Scotland and Iceland, and this may be associated with poor receiver coverage.

The thin shell errors are shown in Figure 5.8 (b). Here, there are large positive errors on the eastern side of the image region, rising to 10 TECu in places. This is equivalent to about 1.6 m at the GPS L1 frequency. The errors reduce as the western side is

approached, and are slightly better than the MIDAS errors in this half of the image region.

5.2.4 Vertical TEC - 28 March 2001 at 18:00 UT

The final example shows the sunset ionosphere for 28th March 2001 at 18:00 UT. From Table 5.1 the Kp index is now 4. Figure 5.9 (a) shows the simulated ionosphere, in which the regions of highest TEC (about 50 TECu) can be seen in the west and south as expected. A smooth east-west horizontal TEC gradient can be observed, with TEC values falling to about 20 TECu in the extreme east. The input plot, Figure 5.9 (b), shows the measured vertical TEC computed from bias-corrected TEC measurements. This agrees with the model, since TEC values in the order of 50 TECu can be seen in the west and south, and values of about 20 TECu are evident in the east.

The inversions are shown in Figure 5.9 (c) and (d). Here again, the full MIDAS inversion displays good agreement with the model, with the easterly TEC gradient being well reproduced. Some distortion of the TEC contours is evident, but this is minimal. The thin shell has reproduced the model reasonably well only in the south of the imaging region, over North Africa. This may be because the receiver-satellite geometry was favourable at this time. However, the thin shell has again failed to reproduce the main easterly TEC gradient that was evident in the model. TEC in the west of the imaging region, from the Southern Spain northwards, has been seriously underestimated. The region of about 20 TECu that is evident in the far east of the model has been moved northwards and spread out in the thin shell inversion.

To quantify the errors, Figure 5.10 was plotted using the same procedure as in Subsections 5.2.1 to 5.2.3 above. The errors in the full MIDAS inversion (Figure 5.10 (a)) range from about +3 TECu to about -3 TECu. The map reveals negligible errors (green areas) over most of mainland Europe. This is probably due to the good receiver coverage in this region. Small localised departures from the model are evident in the extreme south-east (up to +3 TECu) and the south-west (-3 TECu). The thin shell errors (Figure 5.10 (b)) are much larger, ranging from -15 TECu in the west of the imaging region to +5 TECu in the east. 15 TECu is equivalent to a path length error of 2.5 m on the GPS L1 frequency.

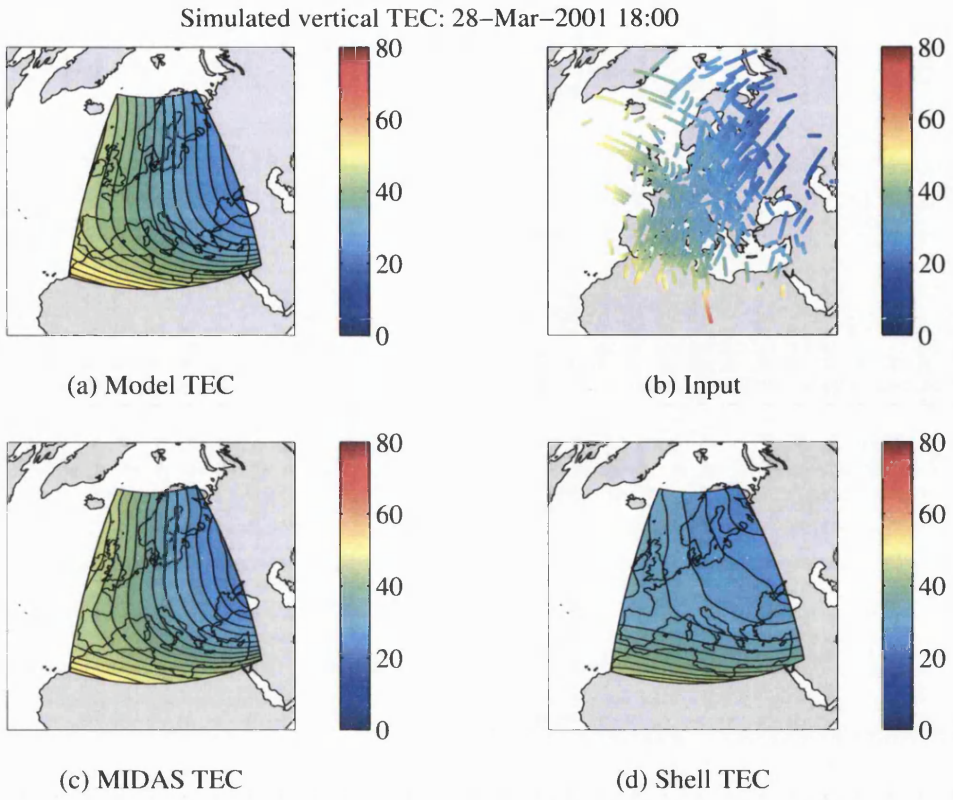


Figure 5.9: TEC maps for 28 March 2001 at 18:00 UT: (a) Model ionosphere from IRI-95, (b) input to reconstructions, (c) vertical TEC map from MIDAS and (d) vertical TEC map from thin shell.

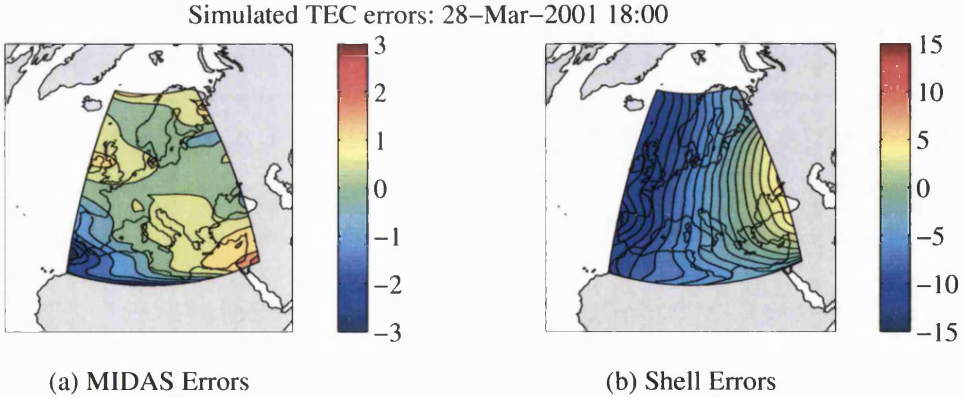


Figure 5.10: Contour plots of absolute errors in TECu between (a) MIDAS inversion and the model, and (b) thin shell and the model.

5.2.5 Mean hourly vTEC errors for all sample days

In Subsections 5.2.1 to 5.2.4, the error in vertical TEC over one hour was examined as a function of latitude and longitude. In this subsection, the rms errors for each hour in the 25-day data set data set are discussed. Denoting the absolute rms error over the entire image region for hour h and day d by $E_{abs}(h,d)$, then

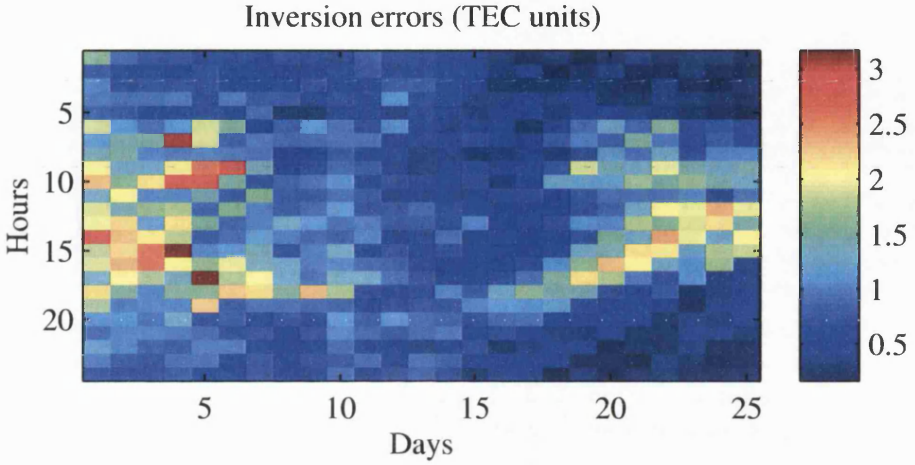
$$E_{abs}(h,d) = \sqrt{\frac{\sum_{i=1}^{N} [I_{inv}(h,d) - I_{mod}(h,d)]^2}{N}} \tag{5.1}$$

Here, I_{inv} is the vertical TEC from the inversion, either thin shell or full MIDAS inversion as required, I_{mod} is the vertical TEC from the model, (h,d) represents dependance upon hour h and day d, and N is the number of latitude/longitude 'squares' in the reconstruction grid. In the present case with 13 longitude increments and 41 latitude increments, N is 533.

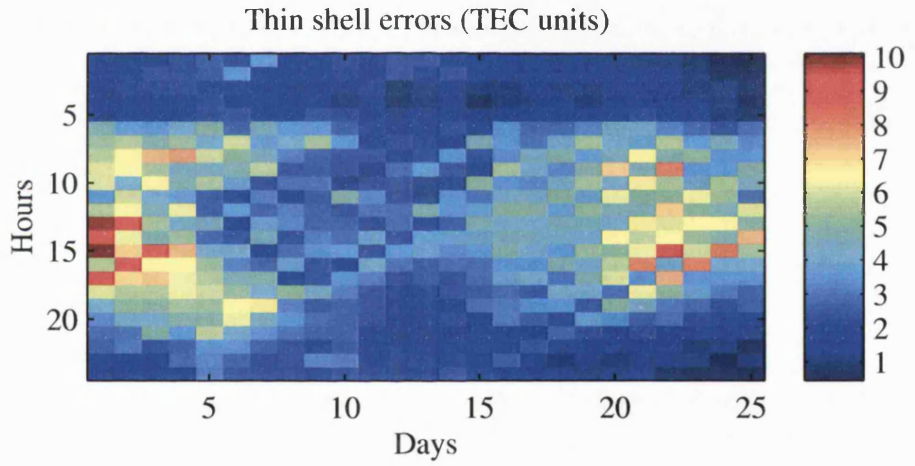
The results were plotted in Figure 5.11, which shows rms vertical TEC errors for full inversion, (a), and thin shell, (b), reconstructions for each hour of the data set. On the x-axis of each figure, the days in the data set are numbered in the order shown in Table 5.1, so that day 1 is 1st January 2001, day 2 is 24th January 2001, etc. The hours in each day are shown on the y-axis, with 01:00 UT at the top and 24:00 UT at the bottom. The negative y-direction reveals the diurnal pattern of the absolute TEC errors can be seen in both plots. Similarly, the seasonal pattern of the absolute TEC errors can be seen in the positive x-direction, with larger errors evident in the winter due to elevated ionisation associated with the winter anomaly. For the whole year, both the full inversion and the thin shell generally reveal rms errors of less than 1 TECu before 06:00 UT, and after 19:00 UT. However, there are exceptions to this, notably in the thin shell plot in which errors of 7.5 TECu can be seen as late as 21:00 UT on day 5, which is 16th March - a quiet day with Kp index of 2 or less.

Figure 5.12 shows the relative rms errors, E_{rel}, for full inversion, (a), and thin shell, (b). These have been computed by normalising the differences between the model and the inversion to the model vertical TEC. In this case, Equation 5.1 becomes:

$$E_{rel}(h,d) = \sqrt{\frac{\sum_{i=1}^{N} \left[\frac{I_{inv}(h,d) - I_{mod}(h,d)}{I_{mod}(h,d)}\right]^2}{N}} \tag{5.2}$$

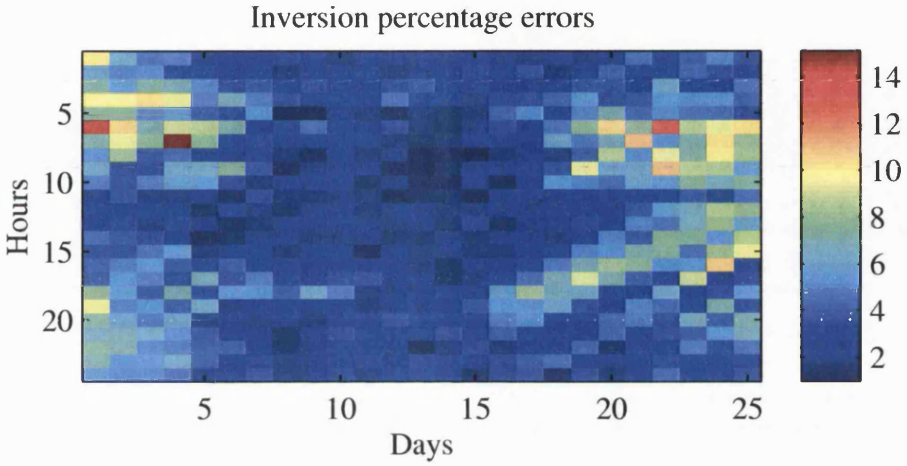


(a)

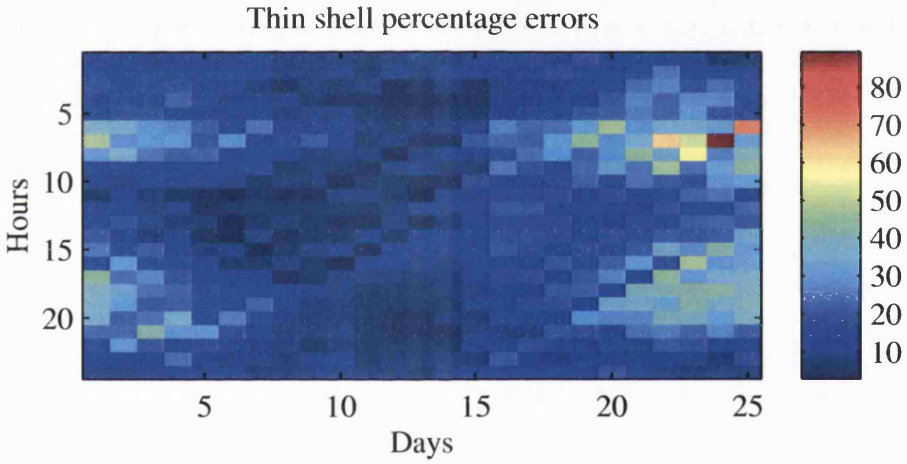


(b)

Figure 5.11: Absolute hourly means of vertical TEC errors for the 25 test days, for (a) full inversion and (b) thin shell.



(a)



(b)

Figure 5.12: Hourly means of vertical TEC errors expressed as percentages for the 25 test days, for (a) full inversion and (b) thin shell.

5.3 Vertical TEC from real measurements

In this section, reconstructions are made for the same times using observations of the real ionosphere. Here, the same satellite dynamics and GPS receiver positions were used as for the simulated case described above, and the slant TEC for all ray paths were then computed from the actual GPS observation data. TEC maps constructed from inversions of real sTEC measurements for each hour in the 25-day data set are presented in Appendix A.

Figure 5.13 shows the results for the night-time of 28^{th} March 2001 at 03:00 UT. In (a), the raw input data (slant TEC measurements) are shown in the form of intersections with a shell at 450 km. These measurements consist of the ionospheric delay plus the satellite and receiver inter-frequency biases. The ionospheric term was isolated by estimating and removing the inter-frequency biases in the manner described in Chapter 3 (see subsection 3.4.3), producing the result shown in (b). Figure 5.13 (c) shows a vertical TEC map produced by the MIDAS inversion using slant TEC measurements through the real ionosphere, and Figure 5.13 (d) shows the TEC map produced by the thin shell method.

It can be seen that these TEC maps bear little resemblance to those in Figure 5.3 (c) and (d), which were generated from the IRI-95 model ionosphere. This is because the model ionosphere is based on a monthly median in geomagnetically quiet conditions. Consequently, small scale structures tend to be smoothed in time and space, and real features that may be present at high latitudes are not visible in the model. Thus the accuracy of the TEC map generated from real data cannot be assessed by reference to the model. However, in the simulations described above good agreement between the MIDAS inversion and the model ionosphere was demonstrated, and since the satellite and receiver geometry is the same in both cases, it is assumed that the MIDAS inversion in Figure 5.3 (c) is a reasonable representation of the real ionosphere.

The TEC map in Figure 5.13 (c) shows vertical TEC values ranging from about 3 TECu in a region covering southern Scandinavia and northern Europe, to about 30 TECu in the extreme south western corner of the imaged region. Steep gradients are evident in the south west and south east corners. Comparing this with the input plot at (b) it can be seen that the overall vTEC distribution is similar. In Figure 5.13 (d), the thin shell reconstruction, the vertical TEC values cover approximately the same range as those of the MIDAS inversion, but the spatial distribution of the TEC is obviously different. In particular, the TEC gradient evident in the south east corner of the MIDAS inversion is absent in the thin shell. The bias-corrected input plot clearly shows slant TEC values approaching 30 TECu in the south east, so it is quite possible that this is a real feature

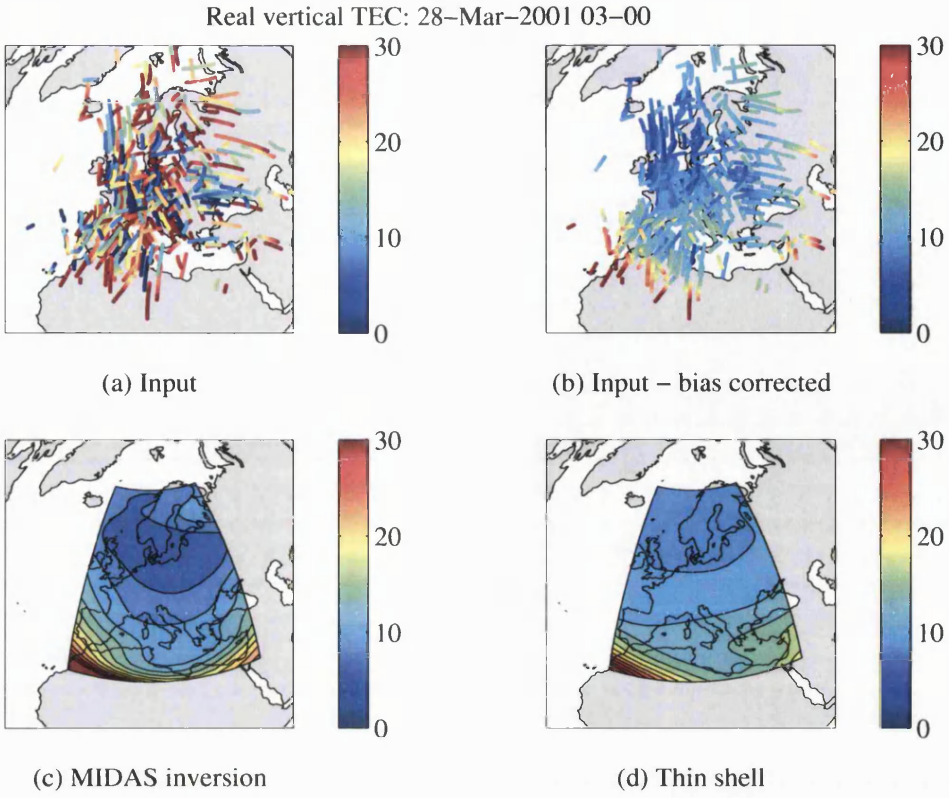


Figure 5.13: Images constructed using real TEC measurements for 28 March 2001 03:00 UT. (a) Raw input plot, (b) bias corrected input plot, (c) MIDAS vertical TEC map, and (d) thin shell vertical TEC map.

that the thin shell has failed to model. The shell has, however, reproduced a steep gradient in the south west, albeit with a different spatial distribution.

The results for the dawn period, 06:00 UT, on 28th March 2001 are shown in Figure 5.14. The raw and the bias-corrected input vTEC plots are shown in Figures 5.14 (a) and (b), and the full inversion and thin shell maps are in Figures 5.14 (c) and (d) respectively. From the bias-corrected input plot, some structure revealing vTEC up to 60 TECu is evident in the extreme south-east, and the remainder of the region displays a maximum vTEC of about 20 TECu. This suggests the presence of a strong horizontal TEC gradient to the south-east, which would be consistent with the early morning gradients associated with the transition from the night-time ionosphere to the daytime. In the full inversion map, a strong gradient is indeed visible in the south-east, and the remainder of the region reveals values of vTEC up to 20 TECu, although vTEC across mainland Europe is slightly lower than would be expected from the bias-corrected input plot. However, the thin shell displays vTEC values over mainland Europe that are closer to those predicted by the input plot, but the gradient in the south-east has not been well reproduced. Indeed, the horizontal gradients that are

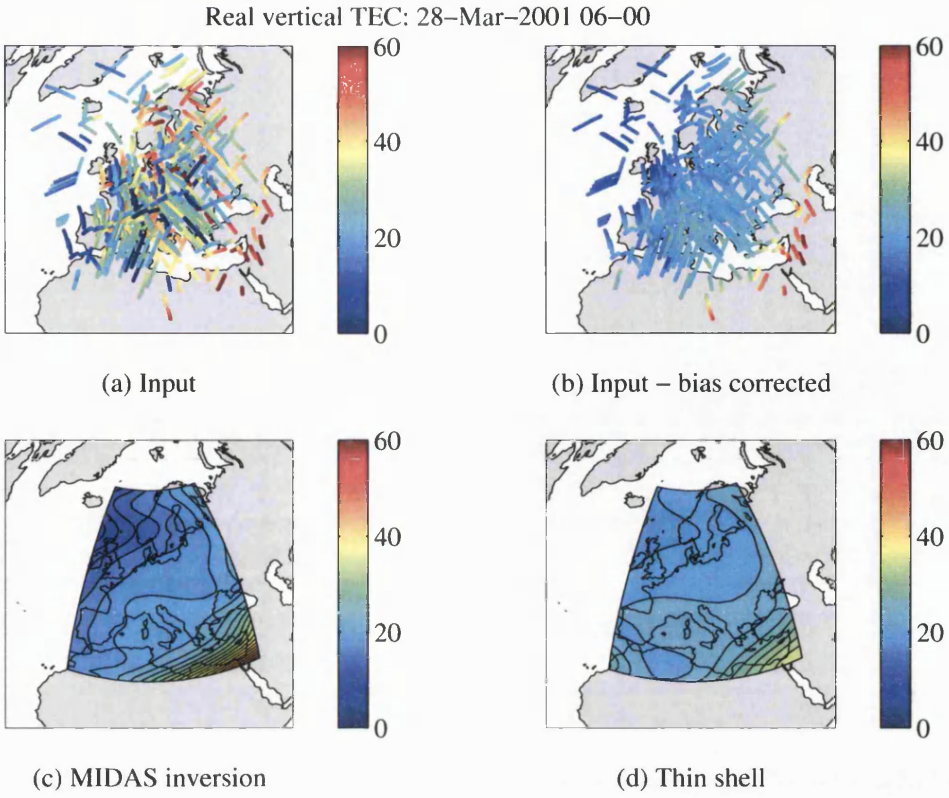


Figure 5.14: Images constructed using real TEC measurements for 28 March 2001 06:00 UT. (a) Raw input plot, (b) bias corrected input plot, (c) MIDAS vertical TEC map, and (d) thin shell vertical TEC map.

evident in the full inversion have, in general, been levelled somewhat.

In Figure 5.15 the images constructed from real TEC measurements for 28th March 2001 at 14:00 UT are shown. Again, the raw and bias corrected input data are shown in (a) and (b). From the bias corrected inputs (b), vertical TEC values of less than 30 TECu (blue) can be expected north west of a line running from the north western corner of Spain to southern Finland. A distinct yellow band of about 50 TECu can also be expected running from southern Spain to the Black Sea. The maps of vertical TEC by MIDAS and thin shell are shown in (c) and (d). Since the MIDAS images in the simulation study produced smaller errors than the thin shell, and the satellite to receiver geometry is unchanged, the MIDAS image produced from real TEC measurements is assumed to be closer to the real ionosphere. Comparing the TEC maps with the bias corrected input plot in (b), two distinct regions of low TEC values (less than 30 TECu) are evident in the north and the west of the MIDAS map, but only in the north of the thin shell map. Some structuring is apparent in the north of both TEC maps. This may be be evidence of a daytime trough, which can be seen more clearly in the sequence of images in Figure 5.20. TEC levels in the west of the thin shell map, particularly

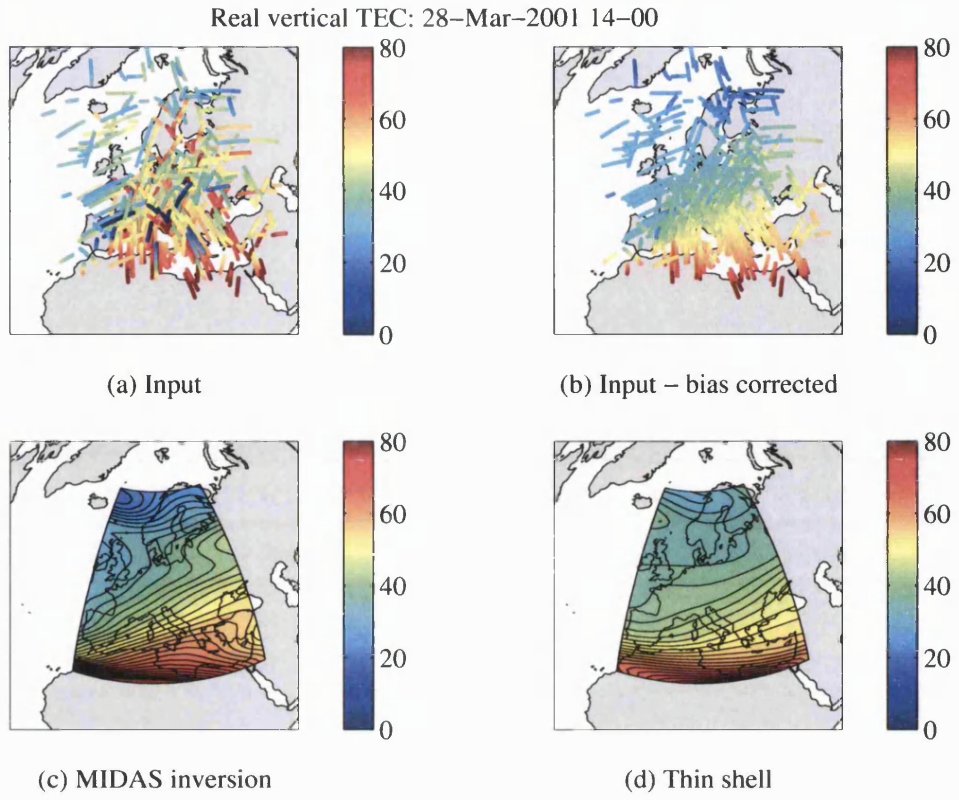


Figure 5.15: Images constructed using real TEC measurements for 28 March 2001 14:00 UT. (a) Raw input plot, (b) bias corrected input plot, (c) MIDAS vertical TEC map, and (d) thin shell vertical TEC map.

over the UK, are higher than in the MIDAS map by about 5 TECu. A distinct band of TEC values of about 50 TECu (yellow) extending from the Black Sea to southern Spain is evident in both maps, but is further south in the thin shell. This has led the thin shell to underestimate the vertical TEC over the eastern Mediterranean. The TEC gradients are generally not as steep in the thin shell map, although there are localised exceptions to this.

The TEC maps for the dusk period, 18:00 UT, on 28th March 2001 are shown in Figure 5.16. The bias-corrected input plot in Figure 5.16 (b) shows evidence of a strong horizontal TEC gradient in the south, ranging in value from 80 TECu down to about 40 TECu. This gradient is reproduced in both the full inversion (Figure 5.16 (c)) and the thin shell (Figure 5.16 (d)) TEC maps. Another horizontal gradient is evident in both techniques, crossing the region from east to west over southern Scandinavia and northern Scotland. This may be the beginnings of the mid-latitude night-time trough, the development of which can be seen in Figure 5.20. The gradient appears to be less steep in the thin shell TEC map than in the full inversion.

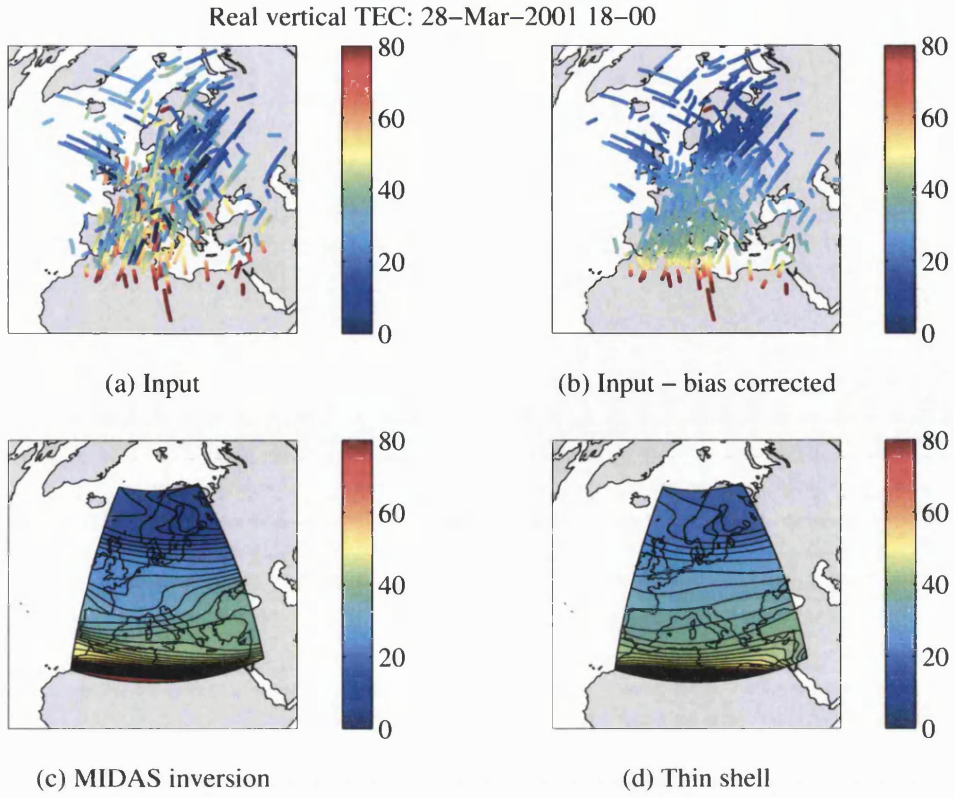


Figure 5.16: Images constructed using real TEC measurements for 28 March 2001 18:00 UT. (a) Raw input plot, (b) bias corrected input plot, (c) MIDAS vertical TEC map, and (d) thin shell vertical TEC map.

In this subsection, the results of inverting real slant TEC measurements for four hours during the day of 28th March 2001 have been discussed. The hours selected represented the night-time, dawn, daytime and dusk ionospheric regimes. Thin shell and full inversion vTEC maps were compared qualitatively with the vTEC at the intersections of the TEC measurements with a 450 km shell.

5.3.1 Hourly maps of vertical TEC

In this subsection, maps of vertical TEC are presented for every hour of 28th March 2001. Figures 5.17 and 5.18 show the horizontal distribution of simulated (i.e., IRI-95) vertical TEC at the start of every hour, for the pre-noon and post-noon periods respectively. As expected, the diurnal variation of the ionisation can be seen rising to a maximum around midday and falling back to a minimum during the night. The night-time TEC values are generally less than 10 TECu, whilst in the daytime TEC values can be seen to rise to a maximum of about 80 TECu. The areas of highest TEC can also be seen to progress in a westerly direction along the southern limit of

the imaging region, confirming the solar dependency of the ionisation processes. The TEC maps also reveal large TEC gradients in the dawn and dusk periods. The dawn gradient starts to appear in the east at about 05:00 UT, and by 09:00 UT has largely passed through the imaging region to the west. A steep TEC gradient is evident in the south of the region from about 10:00 UT until 17:00 UT. The dusk gradients start to appear at about 16:00 UT and persist until 22:00 UT. Over the centre of the region around midday and midnight, the gradients are fairly small, with TEC values of about 50 TECu in the daytime and about 5 TECu at night.

Figures 5.19 and 5.20 show vertical TEC maps at the start of every hour of 28^{th} March 2001, constructed from full inversions of real measurements. Generally, these show low values of TEC during the night-time (up to 10 TECu). However, in the first three hours there is localised structure in the south-west in which the TEC has reached quite high values, up to about 70 TECu in the case of 00:00 UT. Over the three-hour period, the structure can be seen to move to the south-west and decay. Since there are no production processes active at local night-time, this structure is probably a remainder from the previous day. A small area of data in the extreme north-east of the 02:00 UT image is missing. This is believed to be caused by an error in the inversion that resulted in negative values of vertical TEC in that region.

Dawn gradients start to appear at about 07:00 UT, but they are less distinct than those of the simulated ionosphere and merge very quickly into the daytime gradients that are evident along the southern edge of the region. The diurnal pattern of the distribution is evident in these daytime gradients as they move from east to west, eventually dying out altogether around 20:00 UT. The expected dusk gradients are not evident in these images, although it is interesting to note the appearance of a trough in the 19:00 UT image. The trough can be seen to develop over the remaining five hours of the day.

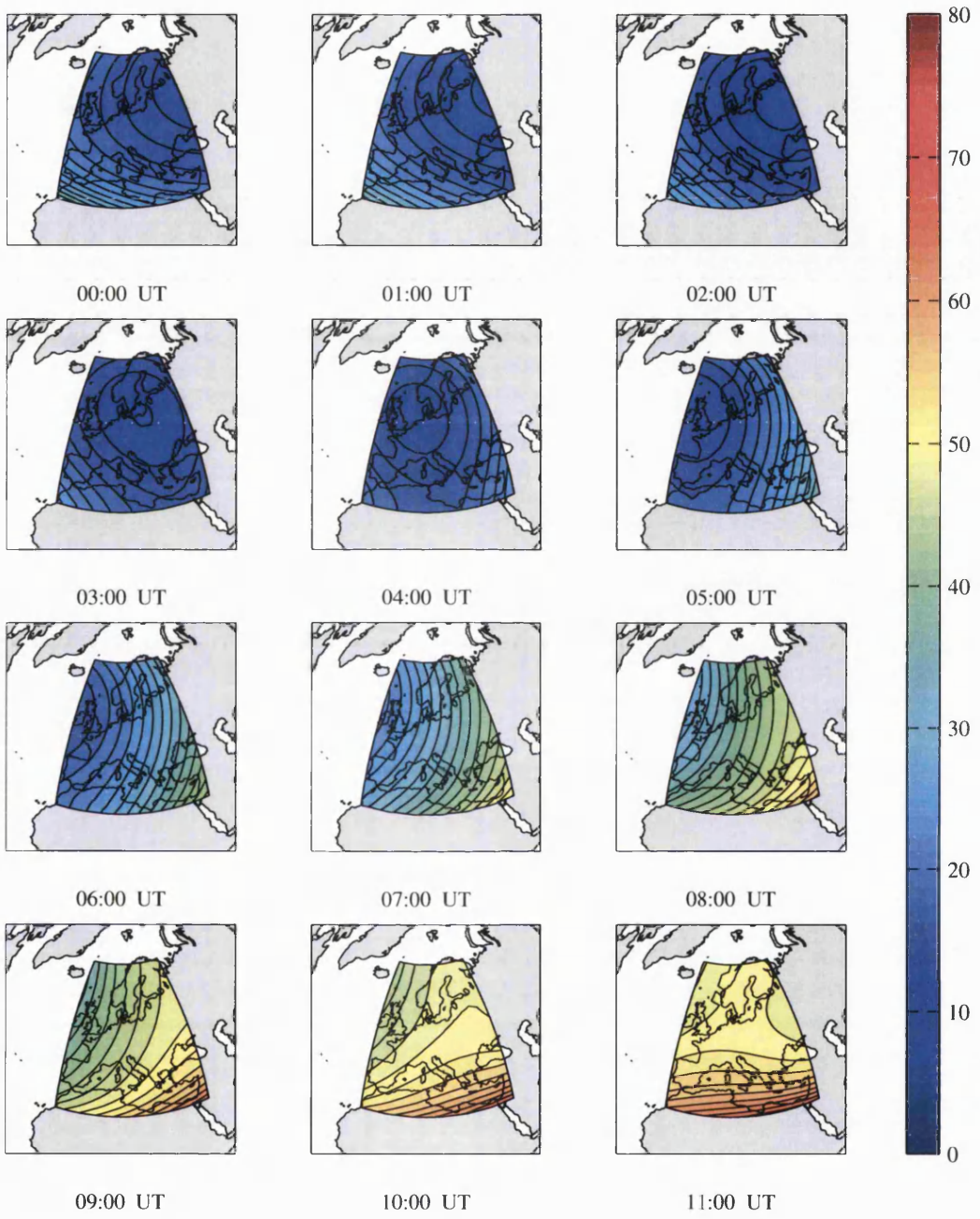


Figure 5.17: Hourly vertical TEC maps of IRI-95 simulated ionosphere. 28 March 2001, 00:00 UT to 11:00 UT.

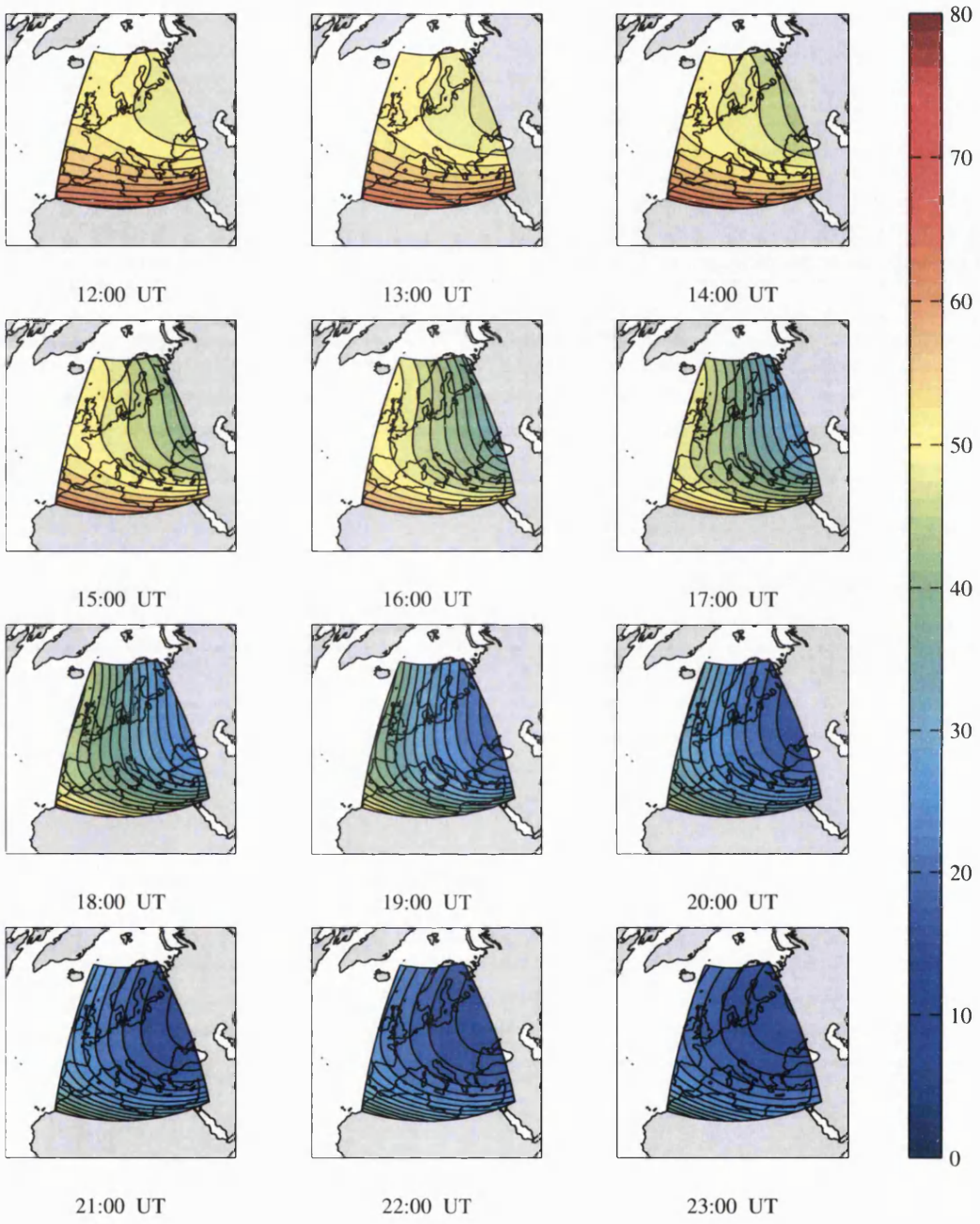


Figure 5.18: Hourly vertical TEC maps of IRI-95 simulated ionosphere. 28 March 2001, 12:00 UT to 23:00 UT.

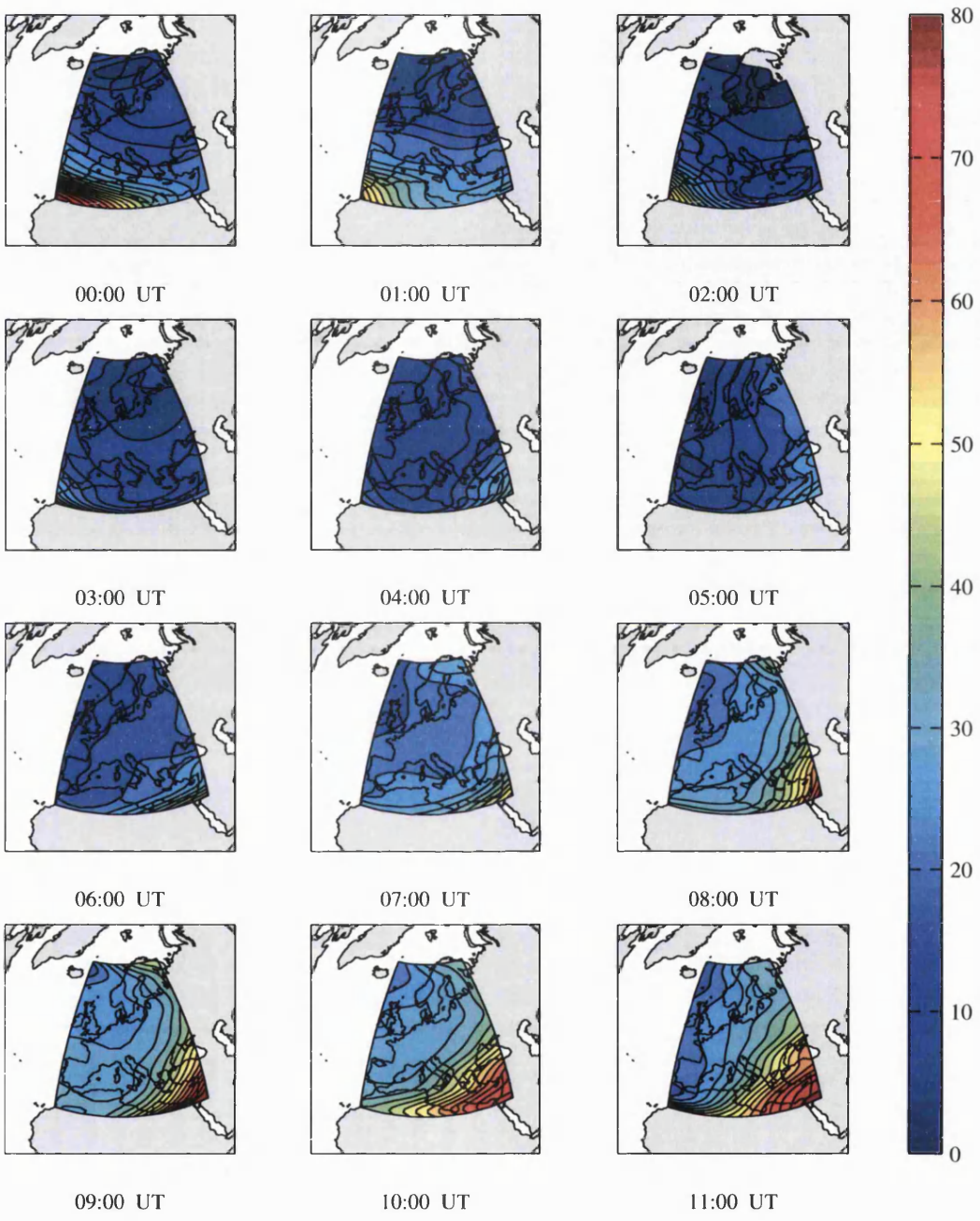


Figure 5.19: Hourly vertical TEC maps from real measurements. Full inversion, 28 March 2001, 00:00 UT to 11:00 UT.

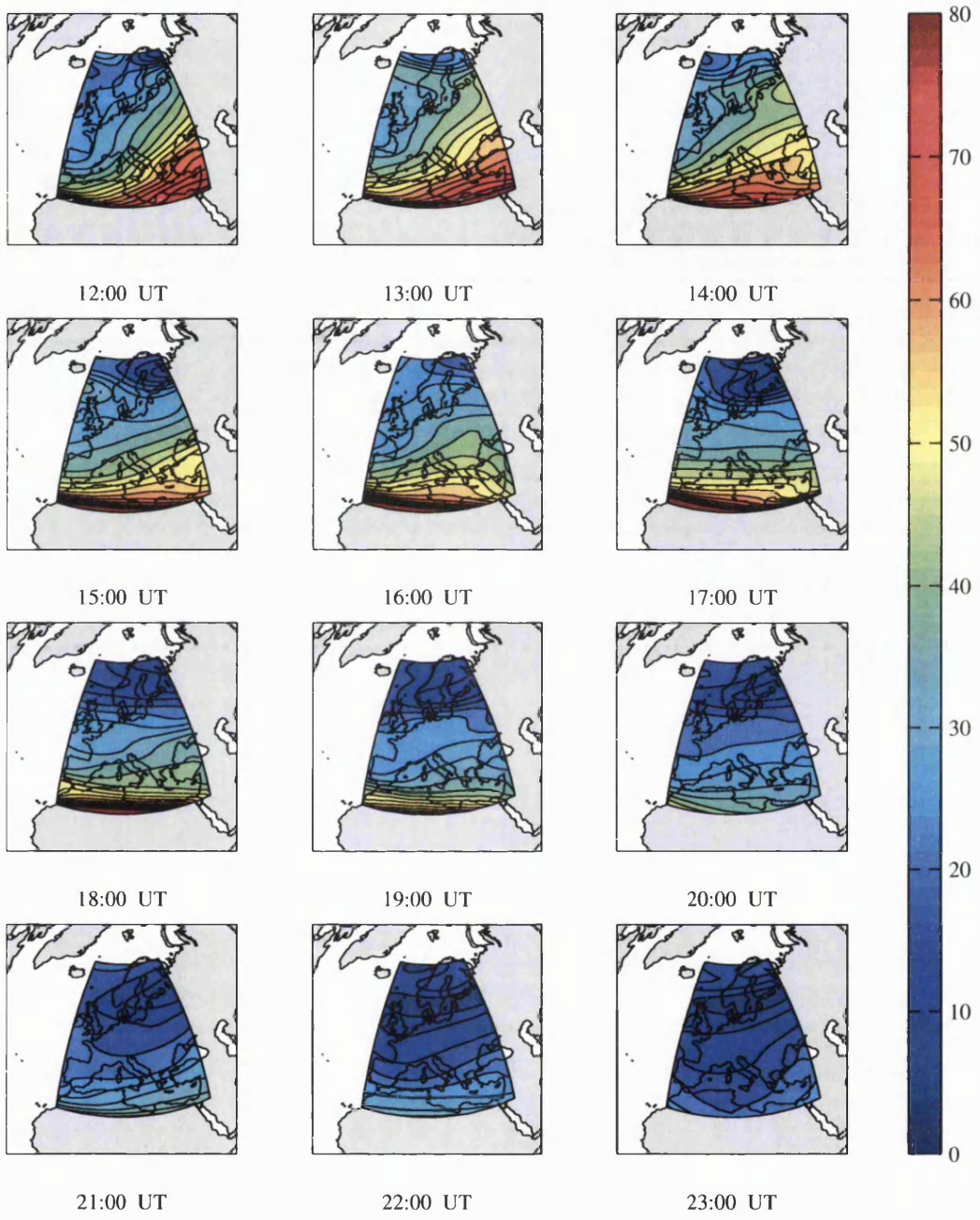


Figure 5.20: Hourly vertical TEC maps from real measurements. Full inversion 28 March 2001 from 12:00 UT to 23:00 UT.

5.4 Slant TEC from real measurements

In this section, the slant TEC errors in the full inversions are examined by comparing each measured slant TEC with the corresponding slant TEC computed through the full inversion. The thin shell technique is not considered in this section, since it is not possible to determine the distribution of electron density along a ray path and hence integrate to obtain the computed slant TEC. Again, the 25 days in 2001 are under consideration. Recall that, from Chapter 2, the inversion problem is expressed by

$$\mathbf{Ax} = \mathbf{b} \tag{5.3}$$

where $\mathbf{A}$ is a matrix of the path lengths in each voxel, $\mathbf{x}$ is a vector containing the electron densities in each voxel of the image, and $\mathbf{b}$ is the measured TEC vector. Assume that $\mathbf{b}$ has been corrected for the receiver and satellite inter-frequency biases as described in Chapter 3. Then it follows that, for a perfect reconstruction of the measured TECs, the residuals would be equal to zero, i.e.,

$$\mathbf{Ax} - \mathbf{b} = 0 \tag{5.4}$$

In general, this is not true; the residuals nearly always have a non-zero value. In the work of this section, this non-zero value is taken as a measure of the error in the slant TEC computed from the inversion.

The procedure adopted here is as follows. The daily mean satellite and receiver hardware biases were estimated for each pair, using the procedure outlined in Chapter 3, and added to the measured paths in the input structures. The resulting set of bias-corrected measured paths was then clipped in latitude and longitude to discard any rays that do not traverse the entire vertical extent of the reconstruction volume. To ensure that only satellites that were above the horizon for the whole hour, any data records with less than 30 measured TEC values were also discarded. The remaining records therefore contain exactly 30 measurements, and all satellite to receiver paths used in the analysis traverse the full vertical extent of the reconstruction grid. To improve computational speed, the 15^{th} ray path (i.e. on the half hour) was extracted, and the remaining values discarded. This left a total of 107732 ray paths in the whole data set, or about 180 ray paths in each of the 600 hours. Finally, using the start and end points of each remaining path measurement, the corresponding TEC, $\mathbf{Ax}$, through the reconstruction grid was then computed.

At this stage there exist two data vectors: the measured TEC, $\mathbf{b}$, and the inversion TEC, $\mathbf{Ax}$. The results are presented as graphs of:

1. measured TEC values in TECu.
2. absolute error, abs $\{\mathbf{Ax} - \mathbf{b}\}$, in TECu.
3. percentage error relative to $\mathbf{b}$, i.e. abs $\{(\mathbf{Ax} - \mathbf{b})/\mathbf{b}\} \times 100\%$.

The whole 25-day data set, containing 107732 ray paths, was considered first, then the different subsets were considered separately. The following subsections discuss the results and present the graphs.

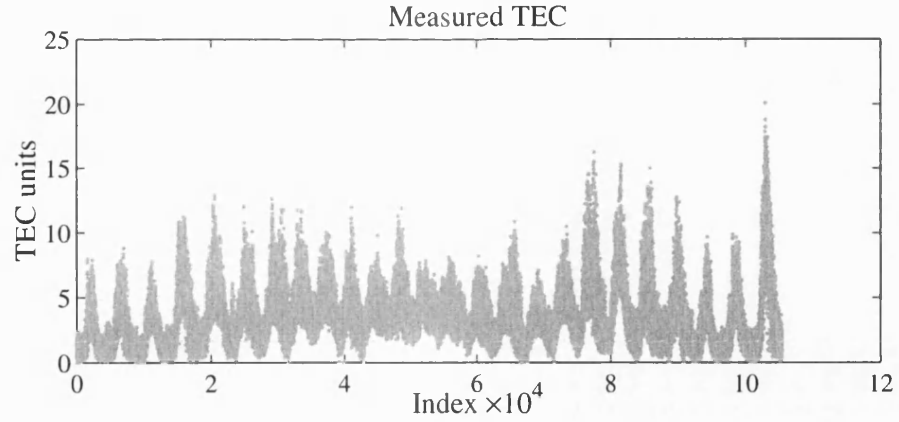
5.4.1 TEC errors for all test days

Figure 5.21(a) shows the measured slant TEC values $\mathbf{b}$ in TECu (1 TECu = 10^{16} electrons m^{-2}) for all 25 days of 2001. Since there are 600 hours in the 25-day data set, each with an average of 180 TEC measurements, the x-axis represents a sequential progression in time of all slant TEC measurements in the data set. Thus each of the 25 peaks visible in the graph represents the diurnal peak in TEC for each day. Counting from the left, the peaks correspond to the day numbers, such that the first peak is day 1 (January 1^{st}) and the second peak is day 2 (January 24^{th}), etc. This applies to all of the results in this section.

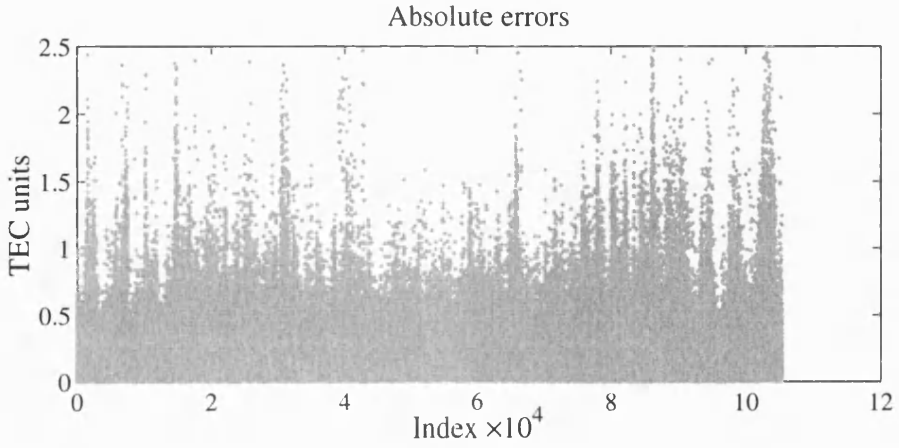
The diurnal peaks can be clearly seen in Figure 5.21(a), ranging from about 8 TECu in January, February, June and July to about 19 TECu in December. The seasonal trend is also discernable, with TEC enhancements evident at the equinoxes. Outlying data points up to a maximum of 50.5 TECu were recorded. The night-time TEC values below about 2 TECu should be treated with caution as TEC values are seldom less than 3 TECu at night (Davies and Hartmann, 1997). Unusually high values of slant TEC can be observed for the last day in the data set, 24^{th} December 2001, which is the subject of separate consideration.

The absolute error in each path was computed by differencing the vectors of measured and computed TEC values, i.e. abs $\{\mathbf{Ax} - \mathbf{b}\}$, and is shown in Figure 5.21(b). Whilst the individual diurnal patterns are less distinct, the seasonal variability is still evident and the enhancements at the equinoxes agree well with those shown in the previous plot. The maximum absolute error was 49.1 TECu, which corresponds to an excess path length of 8 m on the GPS L1 frequency. The mean error, however, was only 0.35 TECu, or about 6 cm at L1, with a standard deviation of 0.68 TECu (11 cm).

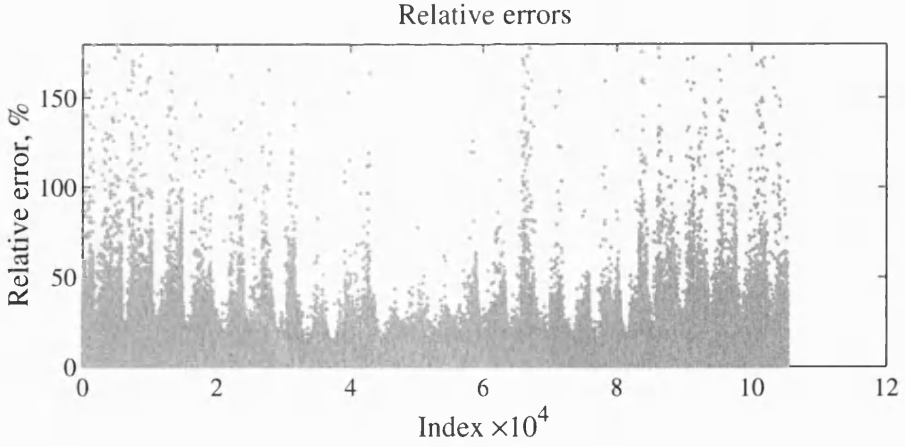
Figure 5.21(c) shows the relative percentage errors, abs $\{(\mathbf{Ax} - \mathbf{b})/\mathbf{b}\}$. Here, the seasonal variability appears as an enhancement in relative errors in winter. This could



(a)



(b)



(c)

Figure 5.21: All measured TEC and errors 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

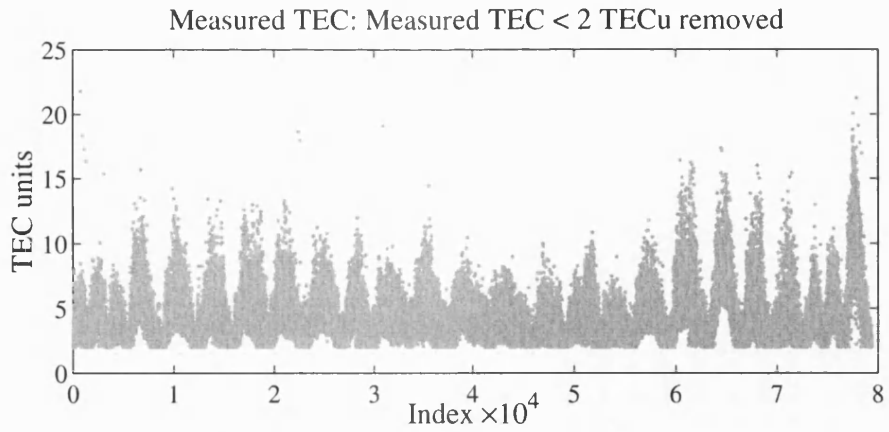
be due to the fact that the nights are longer in winter and so the percentage errors (relative to **b**) will be larger. A few very large values of were recorded for the relative errors, up to 9867%. There were caused by abnormally small values of measured sTEC (**b**). The mean value of this set of relative errors is 14%, with standard deviation of 51%.

5.4.2 Errors after eliminating TEC measurements of less than 2 TECu

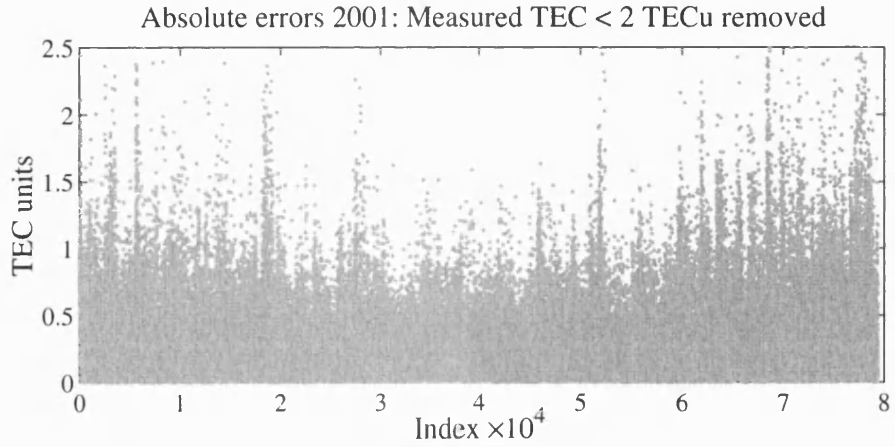
Since the night-time levels of TEC generally reach a minimum of about 3 TECu (Davies and Hartmann, 1997), the results shown in the previous section may have been biased by the inclusion of very low measured TEC values. Here, measured TEC values of 2 TECu or less have been excluded, leaving a total of 79399 ray paths, and the results plotted in Figure 5.22. This still includes any night-time TEC values of around 3 TECu. The measured TEC values greater than 2 TECu for the whole 25 day data set are shown in Figure 5.22(a), which again reveals the TEC enhancements at the equinoxes. The diurnal peaks for the 25 days are also clearly identifiable.

The absolute errors are shown in Figure 5.22(b), and do not show much change in value from those of Figure 5.21(b). There are individual peaks evident for each of the 25 days, but it is difficult to distinguish some of them. The enhancements at the equinoxes are much less evident in this data set, and instead there appears to be an enhancement in absolute errors in the winter-time similar to the situation for the relative errors in Figure 5.21(c). The maximum value of the absolute error in this data set is 49.1 TECu. The mean value is 0.38 TECu, with standard deviation 0.78 TECu.

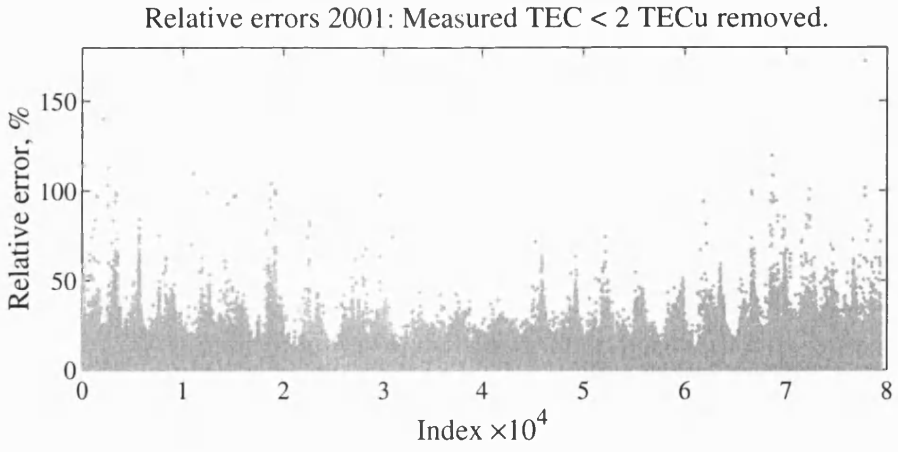
The relative errors for this case are shown in Figure 5.22(c). Comparing this back to Figure 5.21(c) it can be seen that the percentage errors are generally much smaller and less variable. This is to be expected as the TEC values near to zero have been eliminated. The maximum and mean percentage errors in this data set are 172% and 9.5% respectively, and the standard deviation is 9.3%. Also, the underlying trend that is evident in Figure 5.21(c) is less pronounced in this case, and is therefore considered to have been influenced by unusually low TEC values during the night-time.



(a)



(b)



(c)

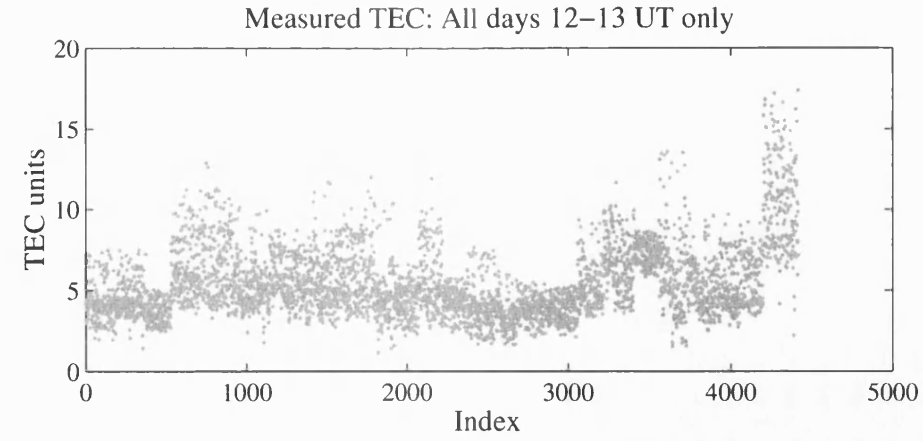
Figure 5.22: Measured TEC > 2 TECu for 25 days in 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

5.4.3 TEC errors for midday on all test days

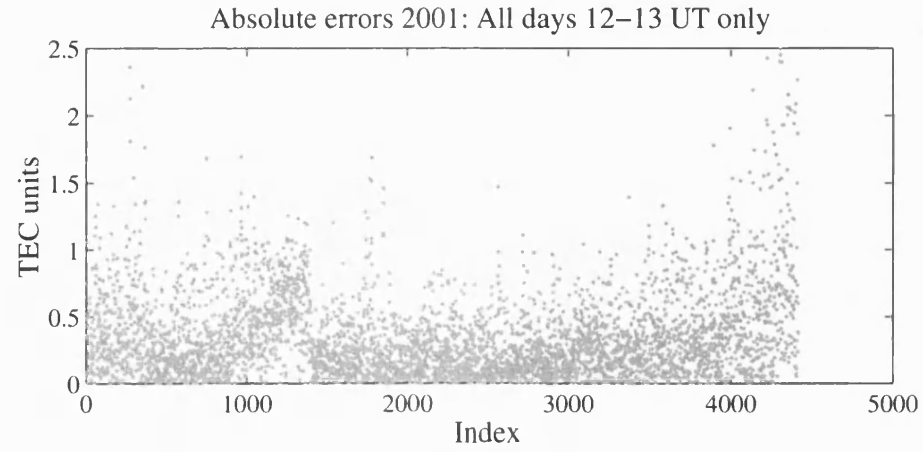
In this subsection, the measured TEC and errors at midday for all the test days are plotted. The purpose of this is to test the possibility that the large percentage errors seen in Figure 5.21(c) are due to the presence of a greater number of low TEC values at winter night-time than at summer night-time. The measured values of TEC are shown in Figure 5.23(a), of which there is a total of 4573 up to a maximum of 28.3 TECu (4.6 m on L1). Again the enhancements at the equinoxes are evident, but the individual days are generally less distinct. This is to be expected as this figure only shows the diurnal peak values.

The absolute errors in Figure 5.23(b) show similar variability to that of the previous figures, with a slight enhancement evident at the equinoxes. The maximum and mean values are 14.3 TECu and 0.4 TECu respectively, corresponding to excess paths on L1 of 2.3 m and 7 cm respectively. The standard deviation is 0.46 TECu. 15 values exceed 2.5 TECu. Of these, 13 lie between 2.5 and 5 TECu, and the remaining two exceed 10 TECu.

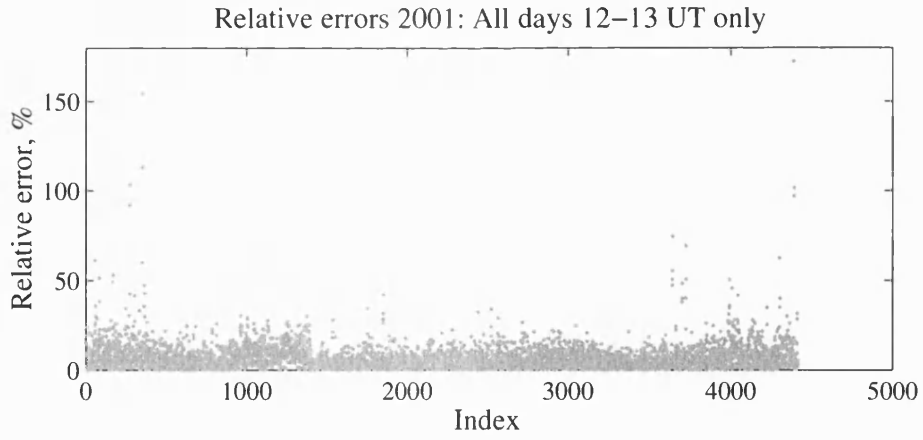
The percentage errors for the midday set are shown in Figure 5.23(c), and reveal a set of values with a mean of 7% and standard deviation of 8%. The underlying trend is negligible. There are about 50 outlying points between 30% (approx 3 sd from the mean) and the maximum of 172%.



(a)



(b)



(c)

Figure 5.23: Measured midday TEC and errors for 25 days in 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

5.4.4 The December 2001 enhancement

The unusually high values of measured slant TEC achieved on 24th December were compared with those of the previous day, 23rd December, and the following day, 25th December. The 3-hourly Kp indices, shown in Figure 5.24, indicate low to moderate geomagnetic activity on the 23rd and the 25th. On the 24th, the geomagnetic activity for the day is moderate, starting with 5 in the dawn period, and falling to 4 in the daytime and at dusk. It is also noted that the solar cycle was near its peak at this time. From Figure 5.25 the monthly mean sunspot number for December 2001 is 115.

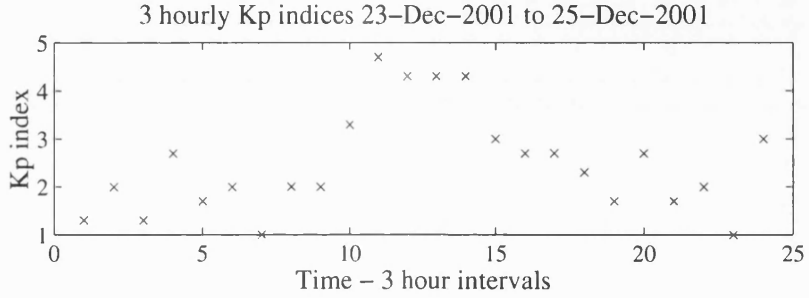


Figure 5.24: 3-hourly Kp indices for the three days of the December 2001 TEC enhancement.

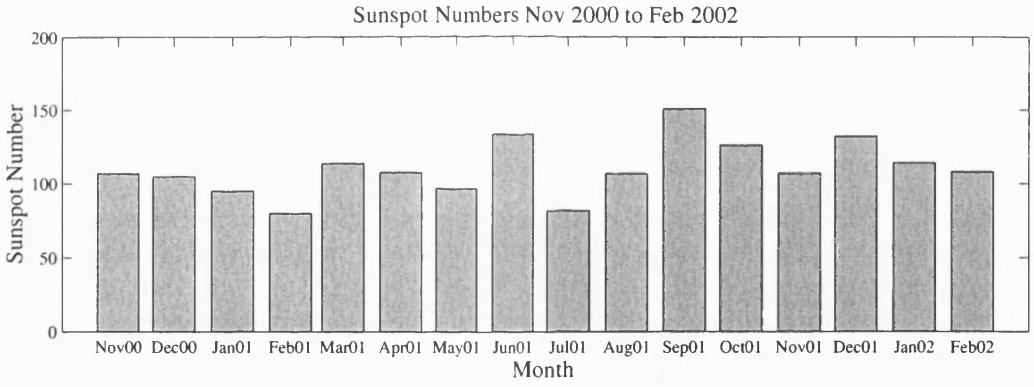
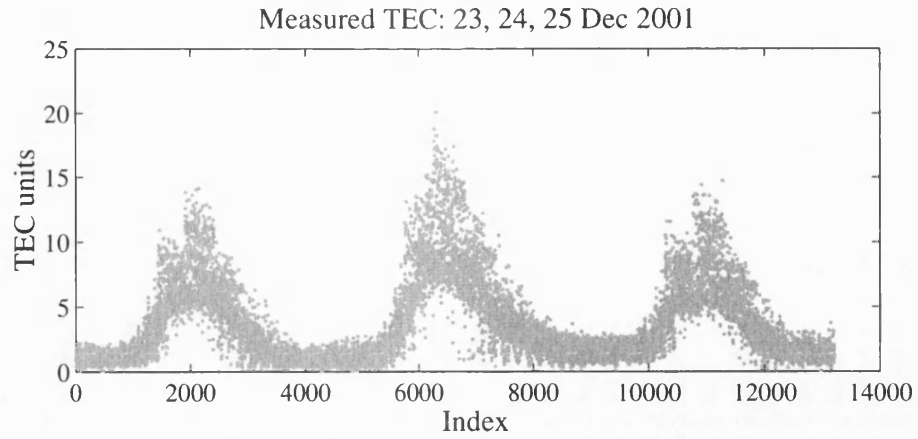


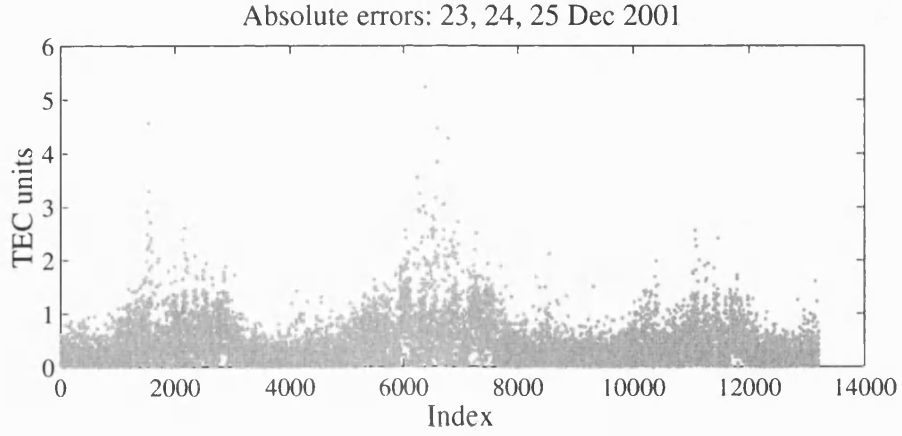
Figure 5.25: Monthly mean sunspot numbers from November 2000 to February 2002. *Source: www.wdc.rl.ac.uk*

A total of 13291 paths for the three days are recorded in Figure 5.26. Comparing the daytime measured slant TEC in Figure 5.26(a), it can be seen that a maximum of about 15 TECu is reached on the 23rd, rising to about 20 TECu on the 24th, and falling back to about 15 TECu on the 25th. Two outlying points of 49.1 and 49.7 TECu can be observed in the night-time between the 24th and the 25th. It is thought that the high values of slant TEC seen here are due to a positive storm effect with

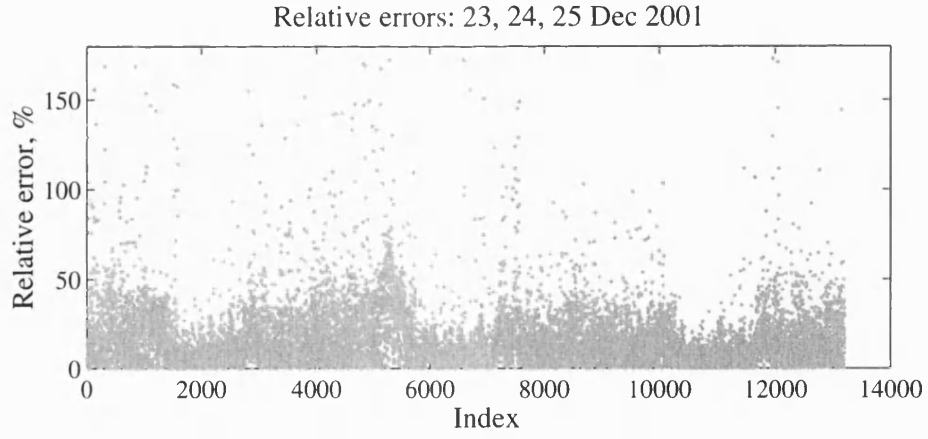
a possible winter anomaly enhancement. The absolute errors for the three days in December 2001 are shown in Figure 5.26(b), and reveal a mean value of 0.44 TECu, with standard deviation of 0.83 TECu. The maximum error, which is clearly biased by the two large values of measured TEC, is 48.2 TECu. The absolute errors can be seen to increase in the daytime. On the other hand, the relative errors, Figure 5.26(c), can be seen to increase in the night-time. They reach a maximum value of 4513%, although the mean value is 18% with standard deviation of 57%.



(a)



(b)



(c)

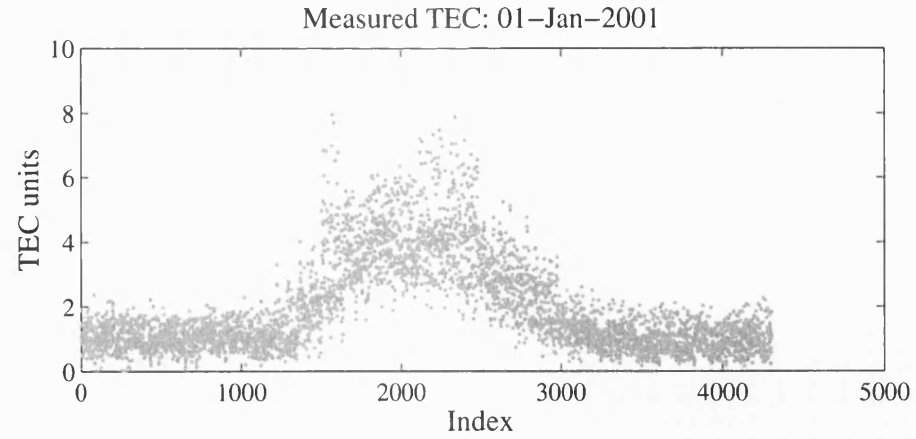
Figure 5.26: Measured midday TEC and errors for 23, 24 and 25 December 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

5.4.5 TEC errors for January

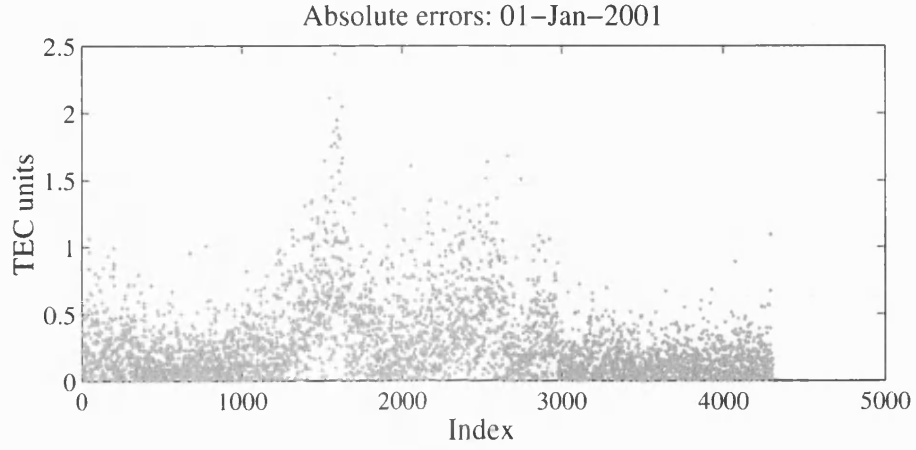
In this subsection, the winter-time slant TEC errors for one geomagnetically quiet day and one disturbed day selected from the 25-day data set are examined. The quiet day is January 1st. From Table 5.1, the 3-hourly Kp indices for that day are between 0 and 1. The disturbed day is January 24th, for which the 3-hourly Kp indices range from 1 to 5.

The results for January 1st are shown in Figure 5.27. A total number of 4328 paths are recorded. The values of measured slant TEC in Figure 5.27(a) clearly reveal the diurnal pattern, which rises from 2 TECu or less in the night-time to a maximum of about 8 TECu in the middle of the day. Four points exceed 10 TECu, with a maximum of 21.8 TECu. The night-time TEC is generally around 2 to 3 TECu. The absolute errors are shown in Figure 5.27(b). They can be seen to rise from 0.5 to 1 TECu during the night-time, to a maximum of about 2.5 TECu in the daytime, although most of the daytime errors are in the order of 1.5 TECu. Four large error values from 13.6 to 14.1 TECu are present. The mean absolute error is 0.3 TECu, with standard deviation of 0.5 TECu. Figure 5.27(c) reveals relative errors that are generally greater in the night-time. The minimum in the relative errors coincides with the daytime values of measured TEC, as revealed in Figure 5.27(a). The maximum percentage error is 52%, and the mean is 22.5%, with standard deviation of 114%.

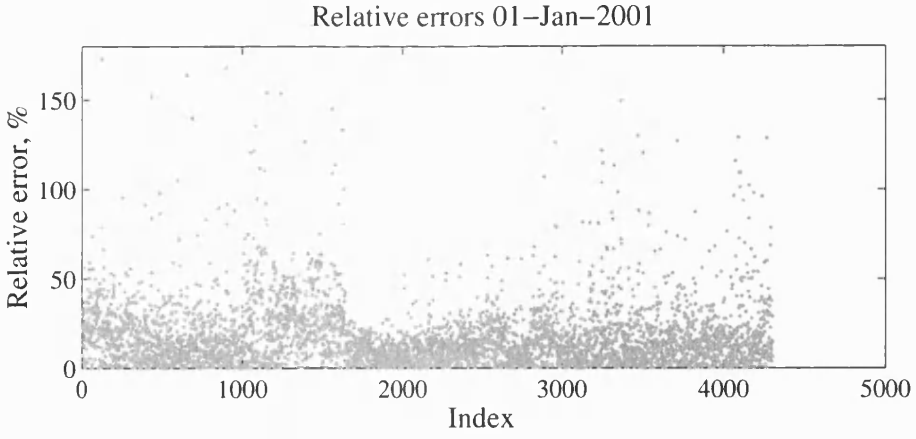
The results for January 24th are shown in Figure 5.28, for which a total of 4702 paths are recorded. Since this is the more disturbed of the two days, it might be expected that values of measured slant TEC would be greater than those of the quiet day. However, this does not appear to be the case. Figure 5.28(a) shows that slant TEC values rise from about 2 TECu in the night-time to about 8 TECu during the day. It is also noticeable that they are more variable throughout the whole day. Four measured TEC with values greater than 10 TECu are recorded, taking the maximum to 49.4 TECu. The absolute errors for this particular day, revealed in Figure 5.28(b), display more variability than those of the quiet day - 1.3 TECu compared to 0.5 TECu. The larger errors are evident in the daytime, as expected. The maximum absolute error is 48.9 TECu, and the mean absolute error is 0.36 TECu. In Figure 5.28(c), the relative errors display two peaks at dawn and dusk. The relative errors during the night-time are again greater than those in the daytime, due to the smaller slant TEC in the night. Maximum percentage error in this case is 1884%. Mean is about 20% with standard deviation about 42%.



(a)

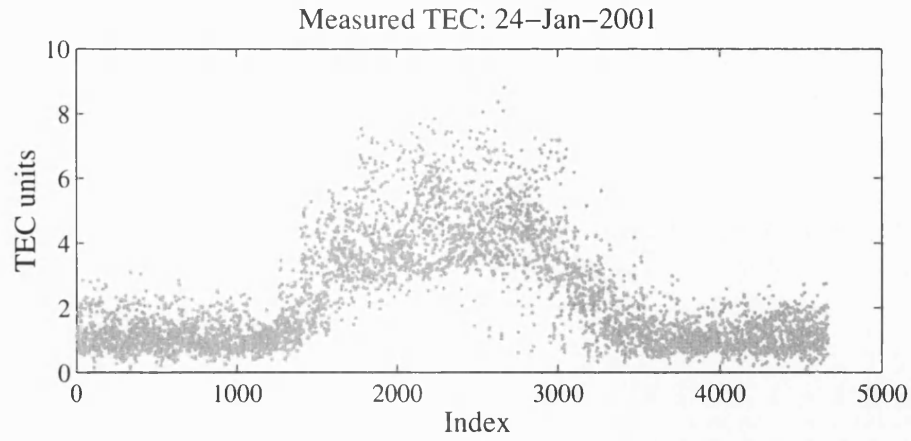


(b)

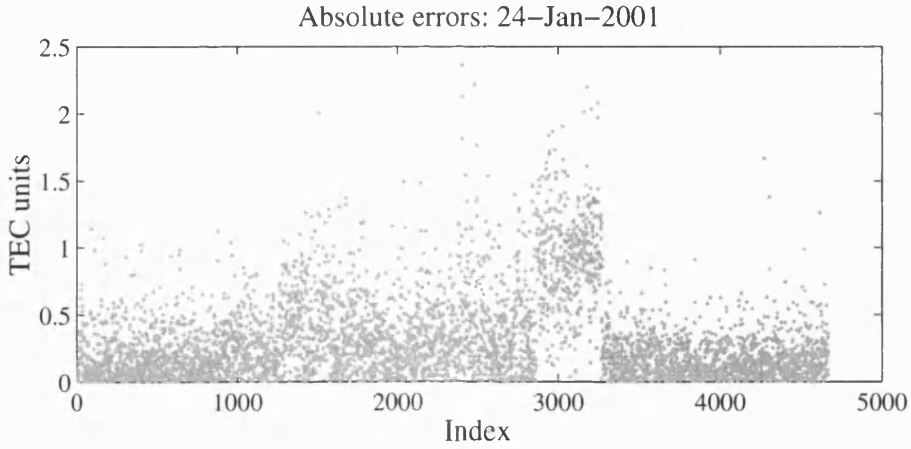


(c)

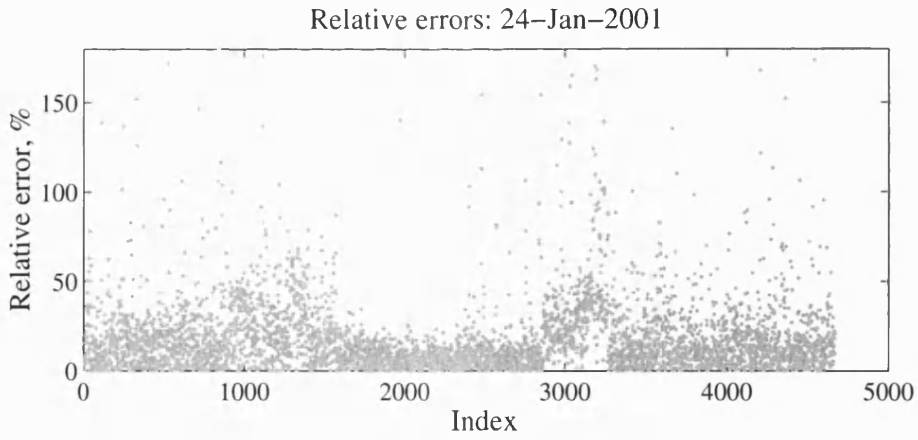
Figure 5.27: Measured TEC and errors for 1 January 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.



(a)



(b)



(c)

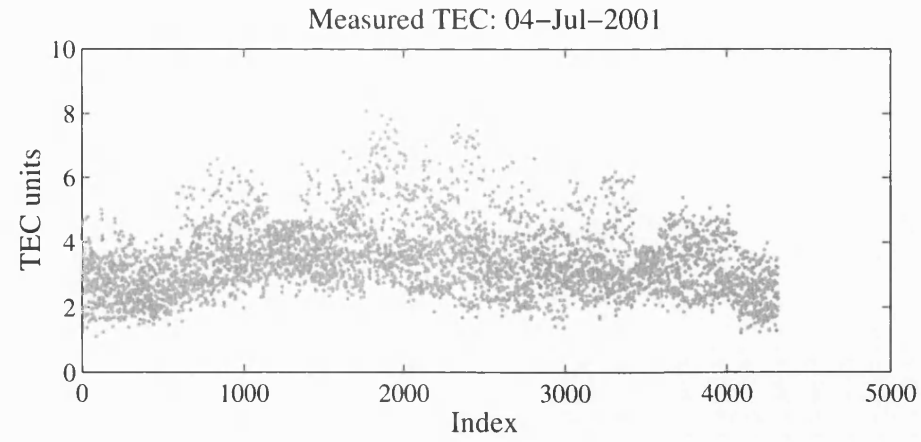
Figure 5.28: Measured TEC and errors for 24 January 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

5.4.6 TEC errors for July

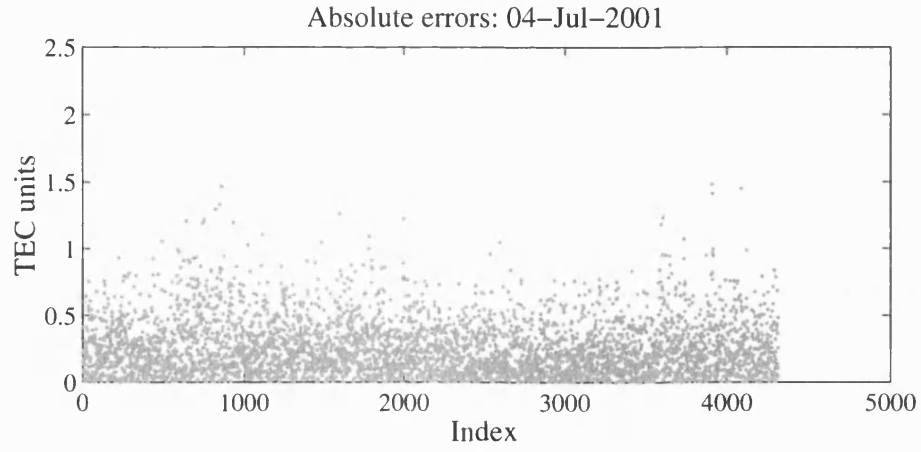
In this subsection, the summer-time slant TEC errors for one geomagnetically quiet day and one disturbed day are examined. The quiet day is July 4^{th} 2001, for which the 3-hourly Kp indices range from 1 to 3 (see Table 5.1). The disturbed day is July 25^{th}, with Kp indices ranging from 3 to 5.

Figure 5.29 shows the results for July 4^{th} 2001. The measured slant TEC in Figure 5.29(a) reveals a minimum slant TEC of just less than 1.5 TECu throughout the day. The maximum daytime TEC rises to about 9 TECu, and the overall shape of the distribution is flatter than those of the winter-time since the daytime is longer. The absolute errors, Figure 5.29(b), are fairly constant with a maximum of about 1.5 TECu, although the mean is 0.26 TECu. This equates to about 4 cm on the GPS L1 frequency. The standard deviation is about 0.2 TECu. The relative errors, revealed in Figure 5.29(c), are quite low (with mean of 8% and standard deviation of 7%) in comparison to those of the winter-time. This is due again to the longer daytime period.

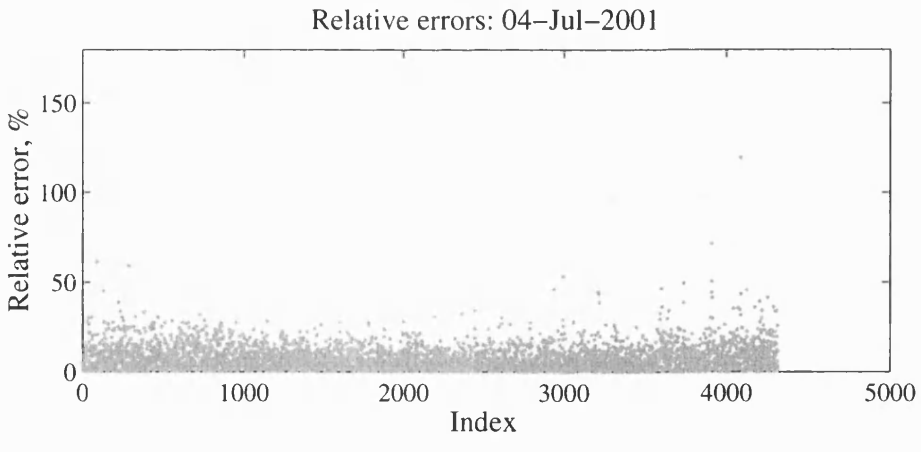
The results for July 25^{th} 2001 are presented in Figure 5.30. The daytime maximum in the measured slant TEC is more clearly defined than that of July 4^{th} (see Figure 5.30(a)), and reaches about 10 TECu. Again, the actual values range between just less than 2 TECu to a maximum of about 8 TECu. A sharp rise in the measured slant TEC is evident in the morning, which could be associated with dawn TEC gradients. Figure 5.30(b) shows the absolute errors. Here, the errors are fairly constant, with a maximum of about 1.5 TECu. A small peak is discernable in the early morning period. The mean absolute error in this case is 0.3 TECu, and the standard deviation is 0.26 TECu. This peak is clearer in the relative errors, and is about 126% (Figure 5.30(b)). They are otherwise reasonably constant with a mean of 11%. The standard deviation is also 11% in this case.



(a)

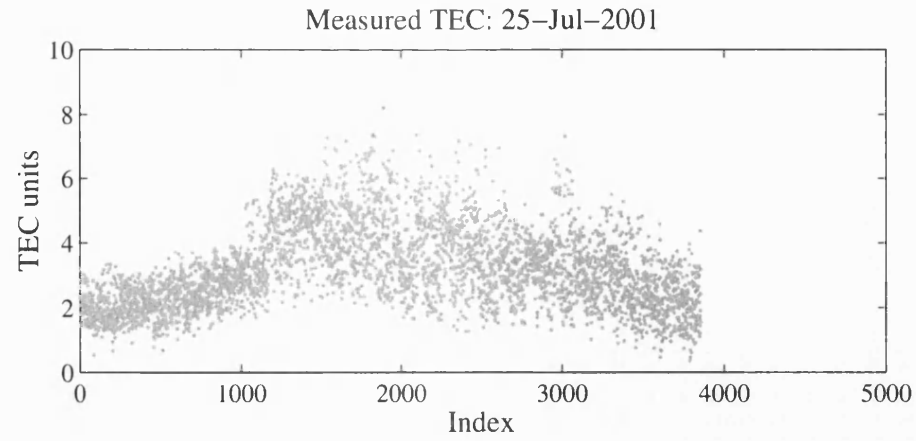


(b)

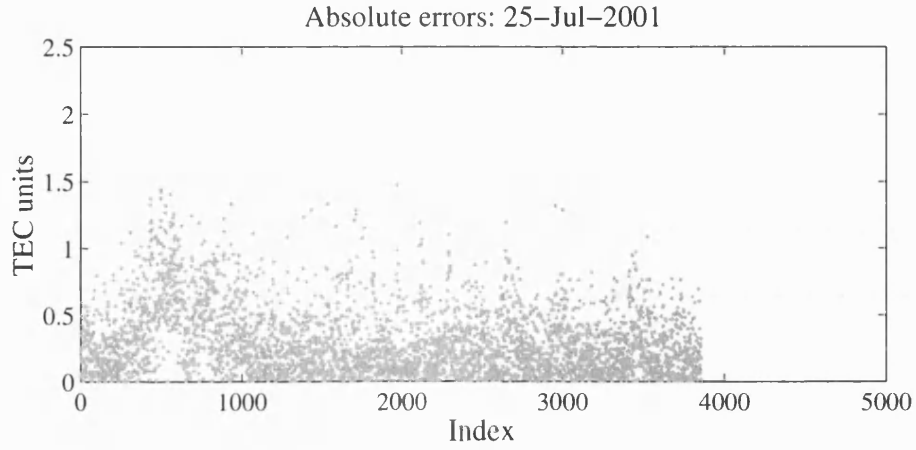


(c)

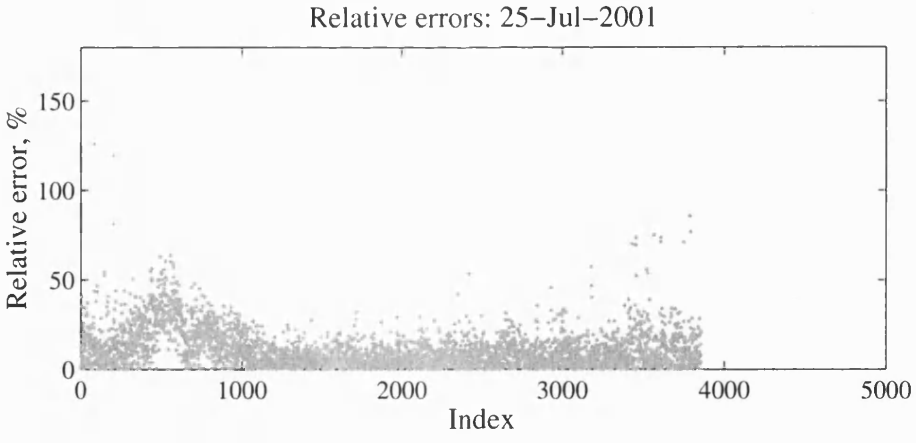
Figure 5.29: Measured TEC and errors for 04 July 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.



(a)



(b)



(c)

Figure 5.30: Measured TEC and errors for 25 July 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

5.5 Discussion

The work presented in this Chapter takes advantage of the recent maximum in the solar cycle, which was reached in 2000/2001 (see Figure 5.31). Since TEC values and GPS signal delays are likely to be at their greatest during solar maximum, this offers an excellent opportunity to test and compare the two imaging methods under 'worst case' conditions. There are three studies presented in this chapter: vertical

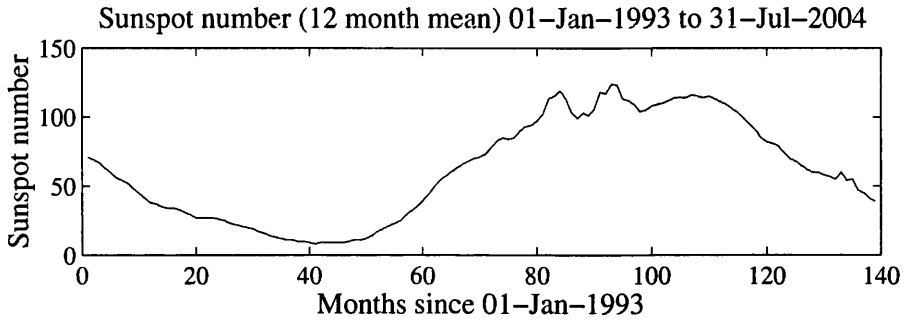


Figure 5.31: Solar activity from January 1993 to July 2004.
Source: www.wdc.rl.ac.uk

TEC from simulations; vertical TEC from real measurements; and slant TEC from real measurements.

In the section on vTEC simulations, the thin shell revealed errors up to around 8 TECu, with no errors higher than 10 TECu. In percentage terms, errors were generally around 35% or less, with maximum errors of about 90%. The inversion showed errors up to about 2.5 TECu, with no errors higher than 3 TECu. In terms of percentages, the errors were generally no higher than 10%, with occasional maximums up to about 15%. In general, a threefold decrease in errors was observed by going from thin shell to full inversion, with greater improvements observed during the day than at night. The thin shell often shows fewer horizontal TEC structures than are present in model, and tends to smooth out the gradients. It is also noteworthy that the higher errors are seen to occur around the terminator, at times associated with strong east-west gradients in electron density. Again, these errors are much greater in the thin shell than the full inversion. This is due to the inability of the thin shell to represent strong horizontal gradients adequately.

In the next section, qualitative comparisons between the thin shell and the full inversion using real data were presented. The simulation study showed that the full inversion could reproduce more real structure from the model than the thin shell. Here also, the full inversion shows more detail than the thin shell. Since this study uses the same

satellite and receiver geometry as in the simulated case, it seems reasonable to assume that the more detailed structure observed in this case is also real. The simulation showed that the thin shell had smoothed out horizontal structure. This was also the case with real data. A good example of this is seen in Figure 5.13, which clearly shows that the thin shell has smoothed out structure in comparison to the full inversion.

The simulation showed vertical TEC errors in the full inversion up to 15%. In contrast, the real data showed estimated mean slant TEC errors of about 14%, but with a very small number of isolated points as large as 9867%. Whilst this would appear to be a very substantial error, it should be remembered that it was computed by normalising the absolute error to the measured TEC. Hence, very small values of measured TEC can be swamped by the TEC obtained from the inversion. The simulation showed lower percentage errors during daytime, and the real data showed lower percentage errors when only 12-13 UT was considered. This implies that the reconstruction is more accurate during the daytime, i.e. when absolute TEC values are greater. However, when considering the absolute error it is found that the night time errors are lower.

5.6 Summary

An extended study has been presented in this Chapter. Observation data from an extensive network of fixed GPS receivers situated in Europe was used to construct images of the ionosphere using both the thin shell and the MIDAS full inversion methods. Twenty-five days in 2001 were selected for this work, which was described in three separate sections. In the first, a simulated ionosphere based on IRI-95 was used as truth, and the vertical TEC obtained by the thin shell and the full inversion techniques was compared. In the second study, a qualitative comparison was made between thin shell and full inversion vTEC maps, which were constructed using the same satellite and receiver geometry as in the first study. Slant TEC measurements were then compared with their corresponding integrated paths through the full inversion grid of electron density in the third study.

Chapter 6

Evaluation of inversion technique using receiver bias stability

Abstract

In this Chapter, a new method to evaluate TEC mapping is presented. The method, described by Meggs *et al.* (2004), makes use of the expectation that the receiver inter-frequency biases are stable on an hourly basis. In the first section, the idea is introduced. Building on the work of Chapter 3, the approach is described in section 6.2. The method is tested by making thin shell and full inversion reconstructions of the ionosphere over Europe at the time of the July 2000 ionospheric storm. Three days ranging in levels of geomagnetic disturbance from 'quiet' to 'very disturbed' were selected. The results are presented for low, mid to high latitude ranges, and demonstrate improvements in the full inversion TEC mapping compared to the thin shell.

6.1 Introduction

In this Chapter, the experimental work is extended by exploiting the stability of the receiver inter-frequency biases as a measure of the slant TEC error in the inversions. Reconstructions of the ionosphere over Europe are made using both the full inversion and the thin shell methods for three days around the ionospheric storm of July 2000. The three days chosen represent a progression of geomagnetic conditions from 'quiet' to 'very disturbed'. In the method described here, selected GPS receivers are grouped in pairs. The pseudo-code and carrier-phase ranging measurements from any pair of

receivers to a common satellite are differenced. This eliminates the satellite inter-frequency bias and leaves only the difference in inter-frequency bias between the two receivers, plus a residual ionospheric-delay difference. Although the absolute value of a receiver's inter-frequency bias is not known *a priori*, it can be expected to remain stable on a short-term basis (see Sardón and Zarraoa (1997) and Mannucci *et al.* (1998)). The standard deviations of the receiver inter-frequency bias differences between pairs of receivers were used as a measure of the relative accuracy of the representation of the ionosphere computed by the two methods. In every case, this criterion showed that the full inversion method gave accuracies in TEC measurement comparable to, or better than, the thin shell method.

To summarise, the procedure used in this Chapter is based on the expectation that the inter-frequency bias of a receiver is constant on an hourly basis throughout the day. Since the absolute value is not known, the difference in the inter-frequency biases of two receivers has been used. This has been obtained by differencing the differential dual-frequency TEC measurements from a pair of receivers to a common satellite, thereby eliminating the satellite inter-frequency bias. The slant TEC error for each satellite to receiver path is then taken to be the difference between the average value of the receiver bias difference and the instantaneous value.

6.2 Method: Bias stability as a measure of error

In Chapter 3, the following expression (Equation 3.25) for a dual-frequency TEC measurement was developed:

$$P'_I = 40.3\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right) I + B_r + B_s$$

Here, P'_I is the ionospheric combination that has been computed by calibrating differential phase-range measurements to differential code-range measurements. Thus P'_I is both smooth and absolute. I is the TEC along the ray path, and B_r and B_s are the receiver and satellite inter-frequency biases. The procedure used here has been described by Meggs *et al.* (2004), and can be summarised as follows. Consider two fixed GPS receivers, k and m, both of which are simultaneously monitoring the same GPS satellite s. Also, for convenience, let $C = 40.3\left(\frac{1}{f_2^2} + \frac{1}{f_1^2}\right)$. Then, rewriting Equation 3.25 for the receivers k and m gives

$$P'_{I,k} = CI_k + B_k + B_s \quad \text{and} \quad P'_{I,m} = CI_m + B_m + B_s \tag{6.1}$$

Differencing Equations 6.1 gives

$$\Delta P_I' = C\,(I_k - I_m) + (B_k - B_m) = C\Delta I_{k,m} + \Delta B_{k,m} \tag{6.2}$$

Thus if the ionosphere is mapped correctly, the inter-frequency bias difference $\Delta B_{k,m}$ will be correctly determined. If the procedure is repeated hour-by-hour, then $\Delta B_{k,m}$ will be constant on an hourly basis. On the other hand, if $\Delta B_{k,m}$ is found to vary then the inversion has incorrectly mapped the ionosphere.

6.3 Slant TEC errors computed from bias stability

Three days before and during the ionospheric storm of July 2000 were selected to test the bias stability method. These were the 11^{th}, 14^{th} and 15^{th} July 2000, representing a progression of geomagnetic conditions from quiet to very disturbed. The 3-hourly Kp values and DST indices for the three days are shown in Table 6.1, from which it can be seen that the third day, 15^{th} July, was the most disturbed of the three. A network of

Table 6.1: Geomagnetic conditions for July 2000 storm days.

Date	3-hourly Kp indices (source: www.wdc.rl.ac.uk)								DST
	03	06	09	12	15	18	21	24	
11 Jul 2000	4	5	4	5	5	4	4	5	12
14 Jul 2000	4	3	4	4	4	6	5	4	-18
15 Jul 2000	4	4	5	5	6	9	9	9	-72

dual-frequency GPS receivers, the locations of which are shown in the map of Figure 6.1, were selected, and the receiver observation and satellite orbital data files obtained from the SOPAC archive at http://sopac.ucsd.edu/.

The differential phase data (smoothly varying, but subject to an unknown cycle ambiguity) for each separate satellite to receiver arc were calibrated to the differential time data (noisy, but useful for calibration) using a least squares fit. Four examples showing the phase data calibration are shown in Figure 6.2, in which the red curves indicate the difference in satellite to receiver path length computed by differencing the L1 and L2 pseudo-code path lengths (in metres), and the blue curves represent the path length difference computed by differencing the the carrier-phase path lengths. As these are measured in wavelengths, they have been converted to metres.

Two sets of reconstructions, one set of thin shell and one set of inversion, were then made for each day. The thin shell height was set at 400 km (see Ciraolo and Spalla

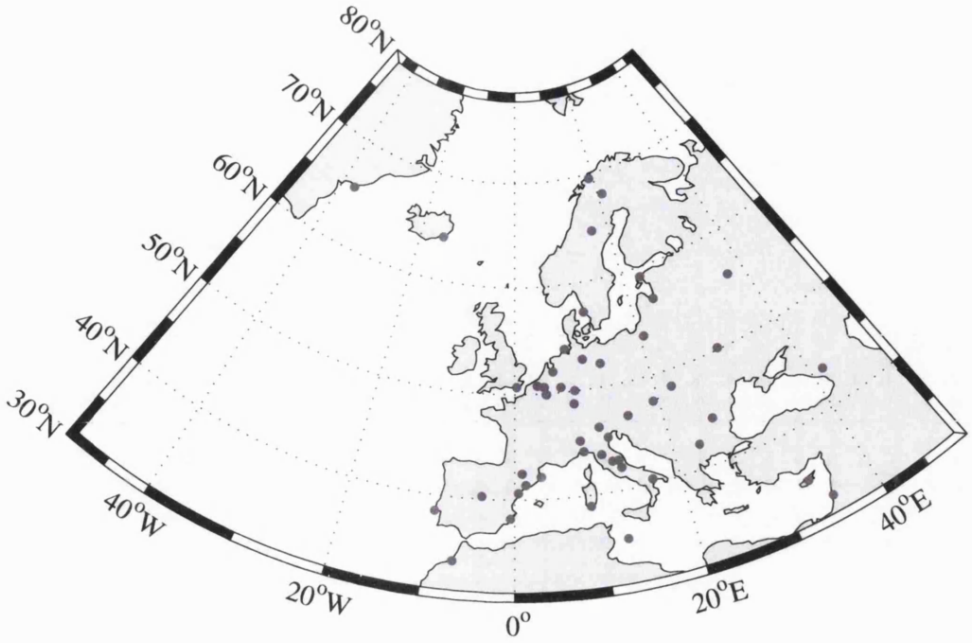


Figure 6.1: Map showing locations of GPS receivers.

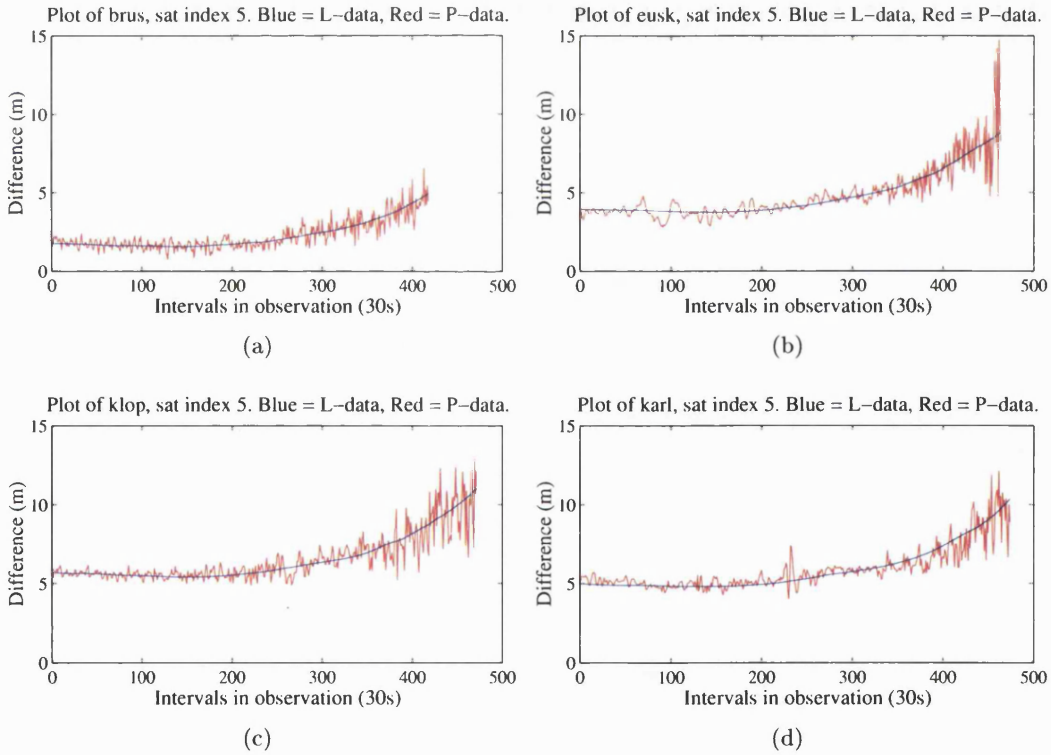
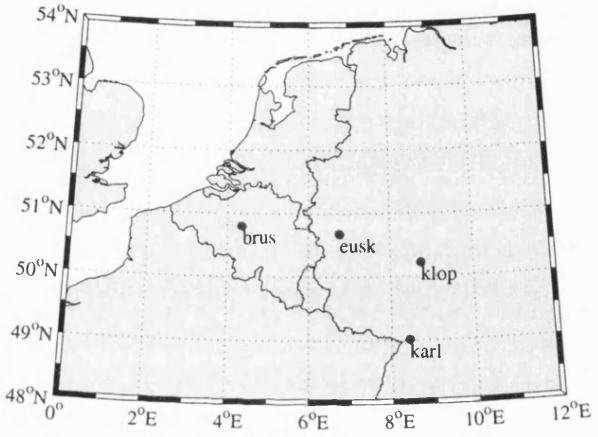
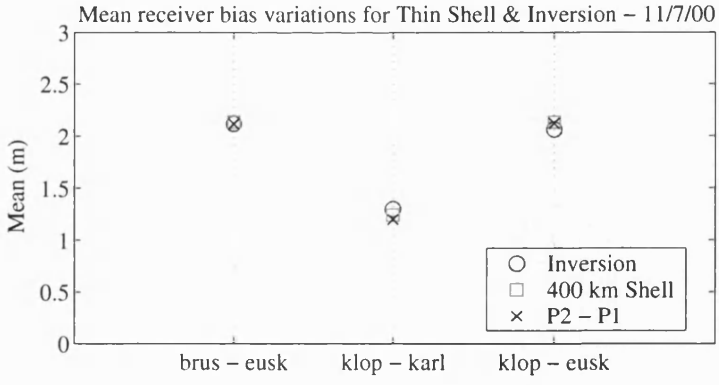


Figure 6.2: Four examples of differential phase calibrated to differential time for 11 July 2000.

(1997) and Chapter 4), and the reconstruction grid for the full inversions method covered a height range from 80 km to 1180 km in steps of 50 km. The horizontal grid spacing used in both methods was set at 1° in latitude and 4° in longitude. To test the computation of the receiver hardware bias differences, the mean hardware bias differences for pairs of close receivers was examined. This provided a set of cases for which the difference in ionospheric delay was negligible. The calibrated phase arcs for a particular satellite in view from two receivers were differenced to provide an estimate of the difference in their receiver hardware biases. A group of four closely spaced mid-latitude receivers situated in Western Europe (see Figure 6.3(a)) were chosen and grouped into three pairs. The mean hardware bias differences for each pair were computed for the inversion and thin shell methods, and plotted in Figure 6.3(b) with the bias differences computed as described in section 6.2.



(a)



(b)

Figure 6.3: (a) Map showing locations of the four receivers and, (b) the mean inter-frequency bias differences for the three methods.

In Figure 6.3(b), the $P2 - P1$ ($\times$) represents the "truth" ionosphere, to which the inversion (circle) and thin shell (square) are compared. It can be seen that all three methods have yielded daily mean hardware bias differences within approximately 10 cm of each other, demonstrating that on a daily average, the inversion and shell are close to the "truth". In each of these specially selected cases, the receiver to satellite paths are very similar. This is to be expected since all paths will experience similar ionospheric delays. Thus, in cases where the ionospheric delay differences can be ignored, the $P2 - P1$ differencing method is a very simple way to determine the hardware bias differences between two receivers.

6.4 Results

6.4.1 Mid-latitude receiver group

The standard deviations of the receiver bias differences for a similar group of mid-latitude receivers (see Figure 6.4) were plotted for all three test days. Figure 6.5 shows

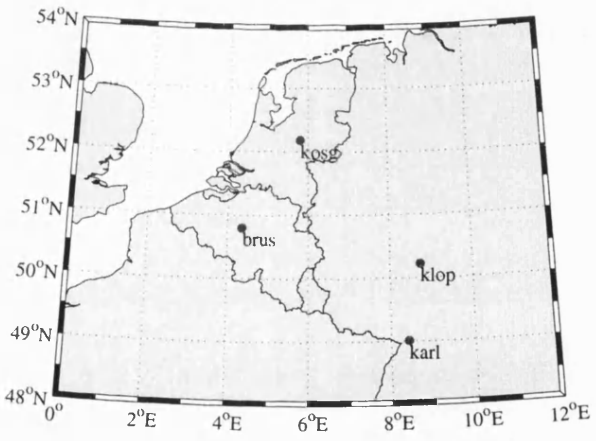


Figure 6.4: Map showing locations and identifers of the group of closely-spaced mid-latitude GPS receivers.

the results, from which it can be seen that there is close correspondence between the stabilities of the inter-frequency biases for both the inversion and the thin shell. This is to be expected, since the receivers are all sufficiently close for their lines-of-sight to any visible GPS satellite to experience similar delays.

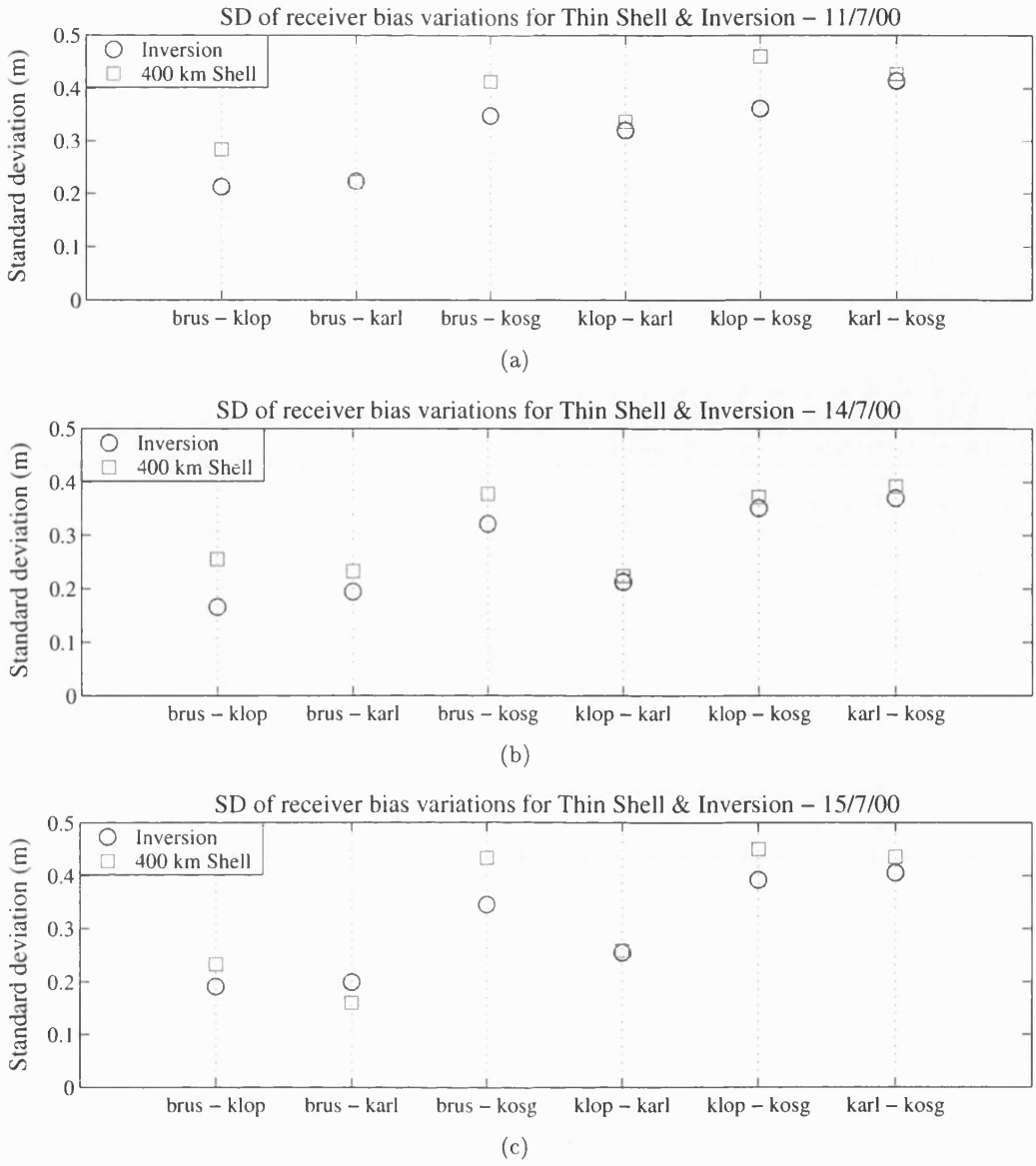


Figure 6.5: Standard deviations of receiver IFBs for closely-spaced receivers: (a) 11 July 2000, (b) 14 July 2000, and (c) 15 July 2000.

6.4.2 Widely-separated receivers

The inter-frequency bias differences for a group of three widely-separated receivers, 'brus', 'onsa' and 'kir0', the locations of which are shown in Figure 6.6, were also examined. Note that the pseudo-code data from 'onsa' reveals evidence of severe multipath effects (see Figure 6.7), and this may have biased the results involving this receiver.

The inter-frequency bias variations for the widely-spaced receiver group are presented in Figure 6.8. In all cases, the inter-frequency bias differences that were computed

from the full inversion exhibited lower variability than those computed using a thin shell. Using the stability of the bias differences as a measure of the accuracy of the inversion, it can be inferred that the inversion method has produced a more accurate representation of ionosphere than the thin shell. As the geomagnetic conditions become more disturbed with the onset of the storm, the inversion standard deviations of the 'brus' - 'onsa' pair can be seen to increase. This may be due to severe TEC gradients existing at the lower latitudes of this group. TEC gradients over Europe may have been significant on 15th July, although the absolute values of TEC did not reach the levels of those over the United States (Mitchell and Spencer, 2002), and hence the absolute errors in TEC mapping would not be expected to have increased.

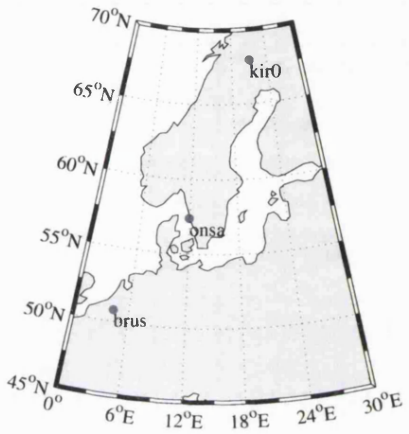


Figure 6.6: Map showing locations and identifers of the group of mid- to high-latitude GPS receivers.

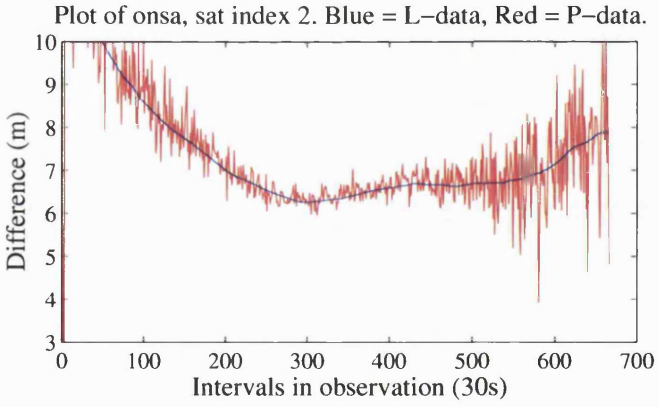


Figure 6.7: Code and carrier phase arcs for the 'onsa' receiver to PRN0002 path showing evidence of severe multipath, particularly towards the end of the satellite pass.

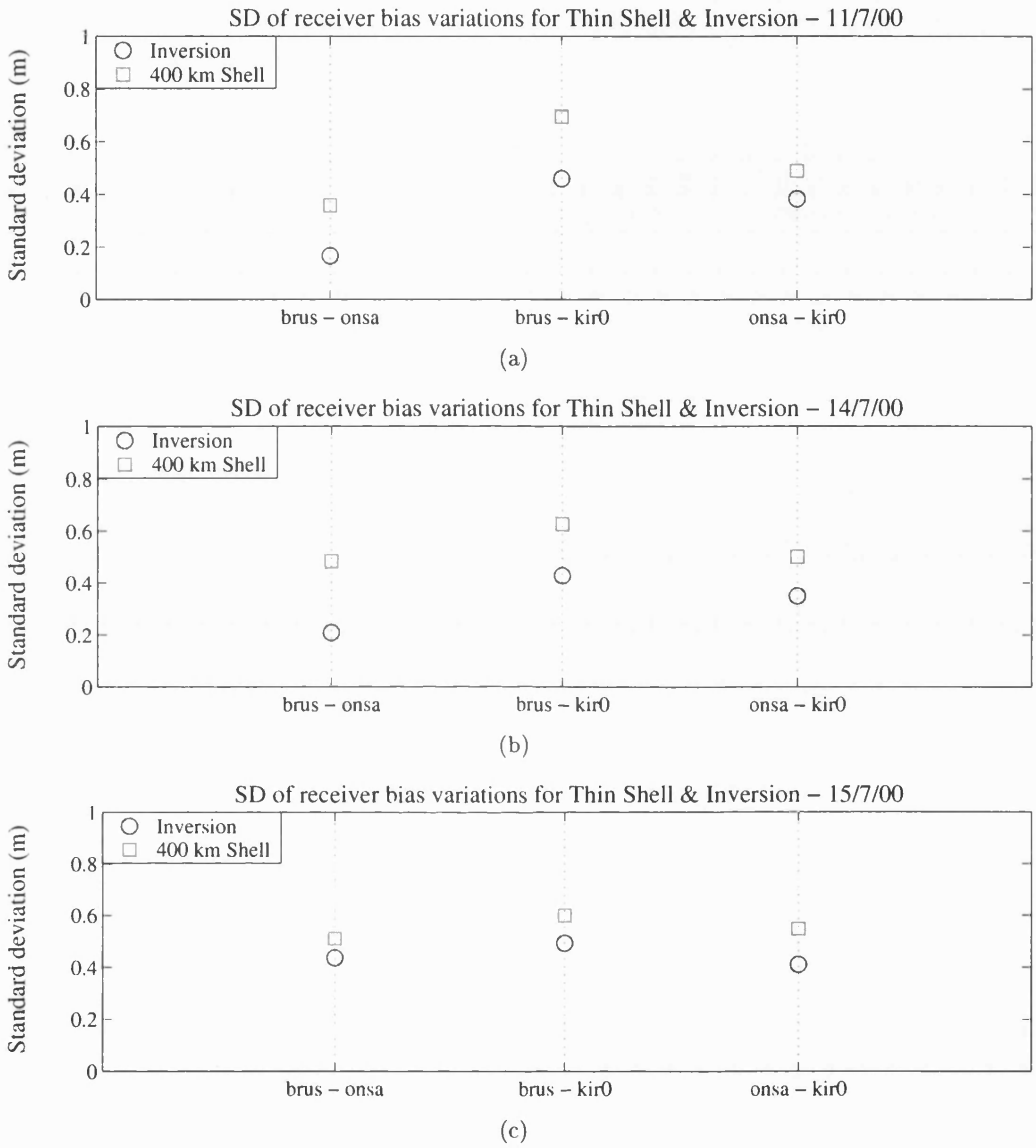


Figure 6.8: All measured TEC and errors 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

6.4.3 Low-latitude receiver group

Figure 6.9 shows the locations of a group of receivers at lower European latitudes (between 35°N and 43°N), and the corresponding standard deviation plots are shown in Figure 6.10. Using the bias stability criterion, the 'lamp' - 'mate' and the 'lamp' - 'aqui' pairs show an improvement in the representation of the ionosphere using the inversion. The 'mate' - 'aqui' pair displays a variability of less than 20 cm for both thin shell and inversion. This implies that the ionosphere is well represented using either method. Since these two receivers are located reasonably close to each other, close correspondence in the results can be expected.

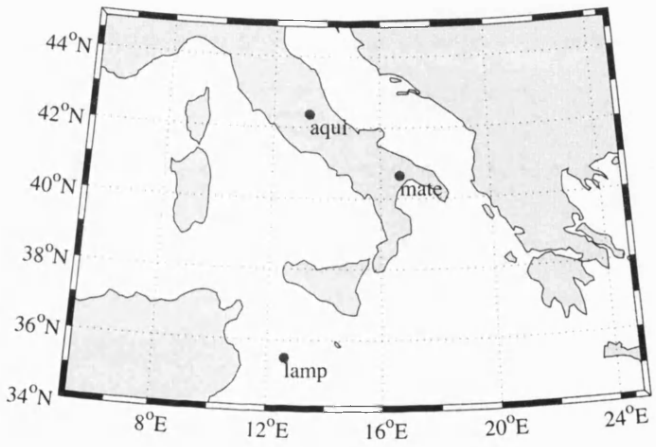


Figure 6.9: Map showing locations and identifers of the group of low-latitude GPS receivers.

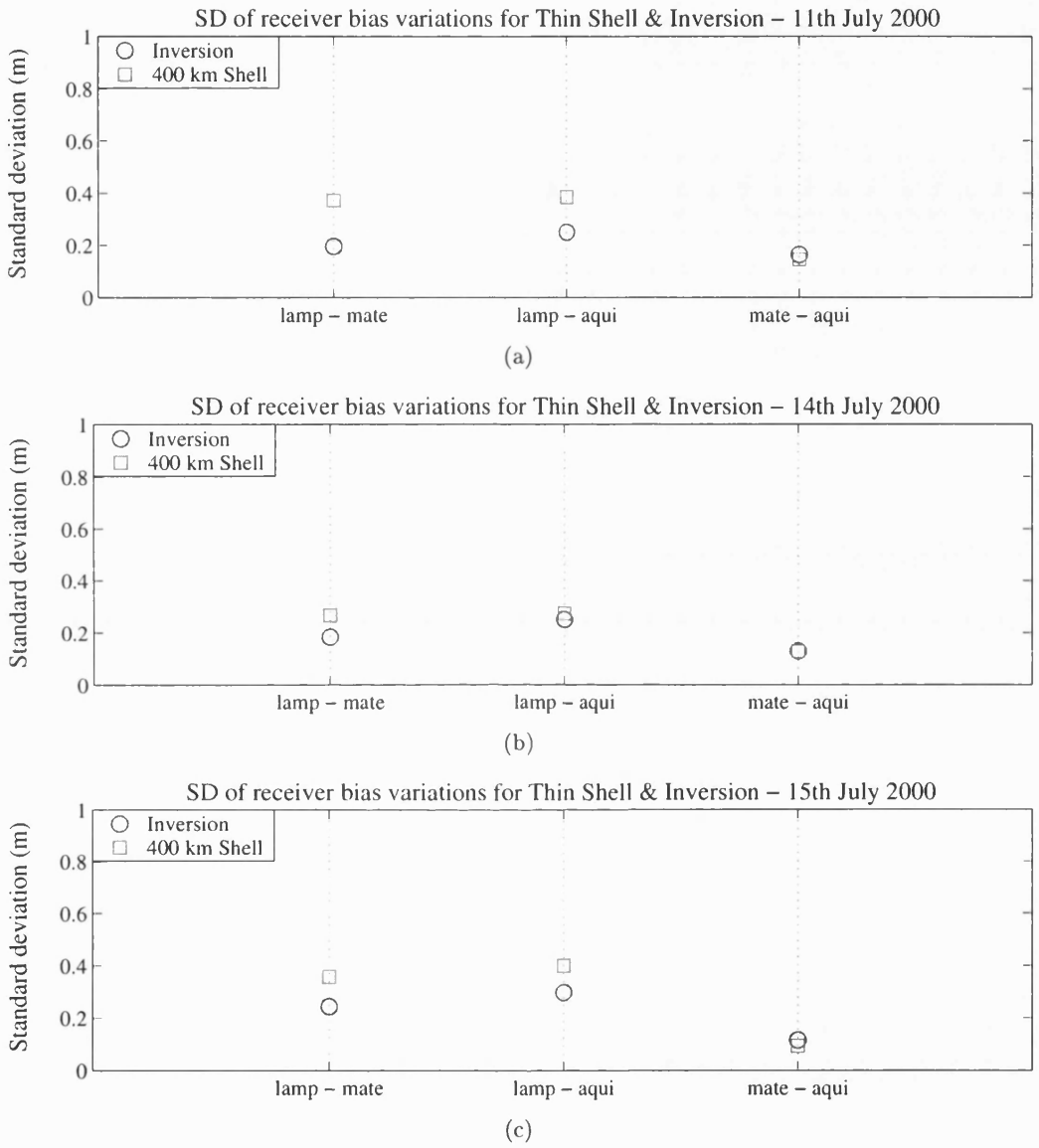


Figure 6.10: All measured TEC and errors 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

6.4.4 Scandinavian receiver group

A receiver group in Scandinavia is shown in the map of Figure 6.11. Again, the standard deviations of the inter-frequency bias differences reveal an improvement in the representation of the ionosphere using the inversion method, particularly with the 'kir0' - 'onsa' pair. This pair are the most widely separated of the three pairs, and the results will be influenced by large horizonal TEC gradients. The 'vil0' - 'kir0' pair display the smallest variability, which increases slightly on the storm day (the 15^{th}). Both of these receivers are at high latitudes (above 64°N).

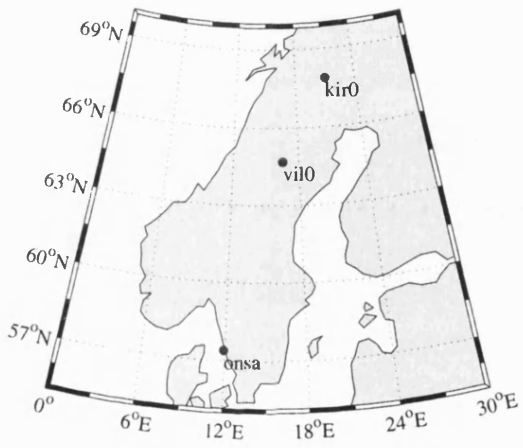


Figure 6.11: Map showing locations and identifers of the group of Scandinavian GPS receivers.

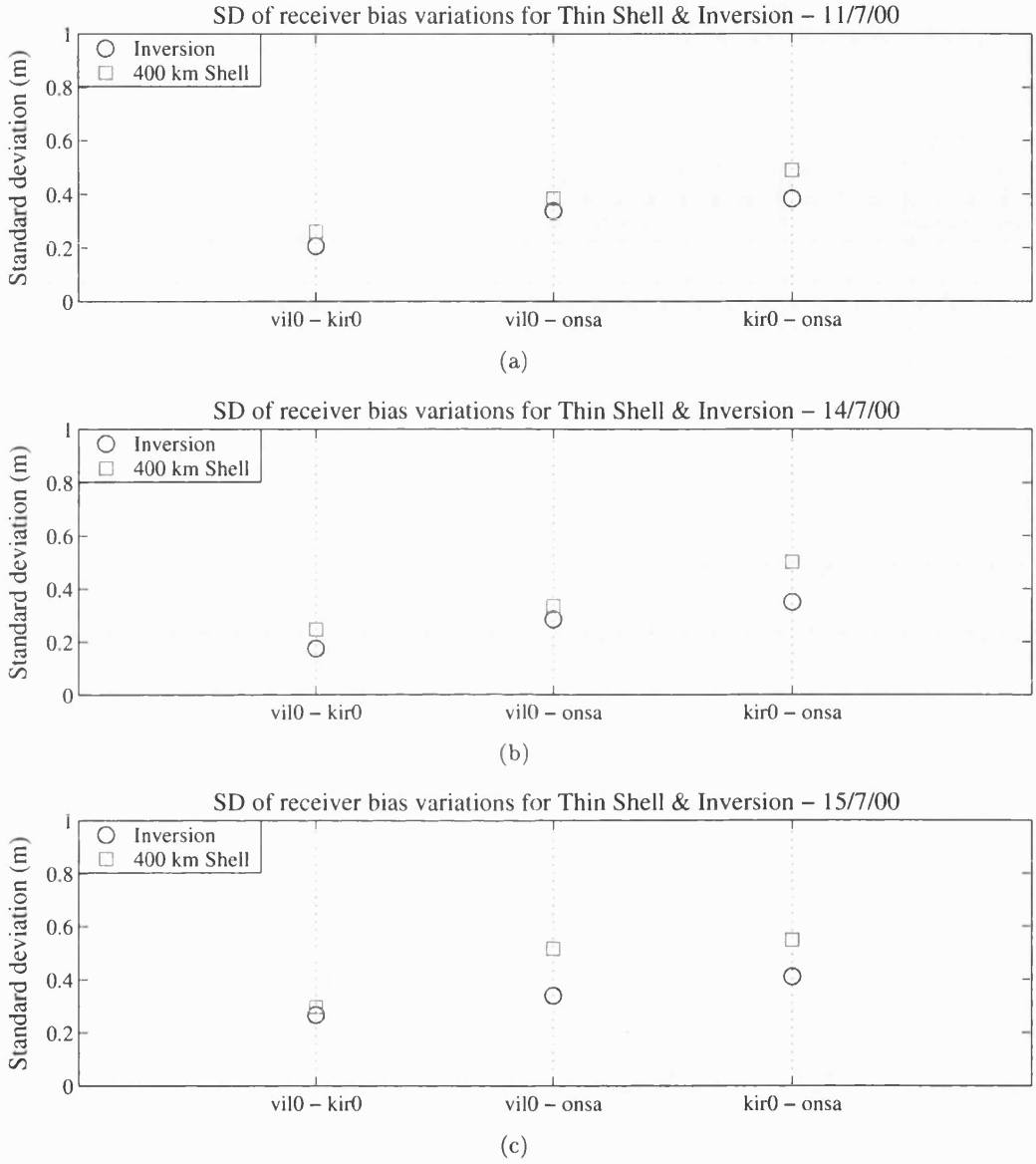


Figure 6.12: All measured TEC and errors 2001: (a) Measured TEC, (b) absolute errors and (c) the relative percentage error.

6.5 Discussion

A new method to evaluate TEC mapping has been presented in this Chapter. In the studies presented in Chapters 4 and the initial part of Chapter 5, which were based on simulated ionospheres, the evaluation of the TEC mapping was a straightforward matter of differencing between the inversion TEC and the simulated TEC. In the experimental case, there is no readily accessible 'truth' ionosphere with which to compare inversions.

The new method relies on the expectation, widely reported in the literature, that the receiver inter-frequency biases are stable on a short-term basis. By differencing the TEC estimates from two receivers to the same satellite, the satellite inter-frequency bias will cancel out and the receiver inter-frequency bias difference will remain. Provided the ionosphere is correctly mapped, the receiver inter-frequency bias difference should be constant. Conversely, if there is a change in the inter-frequency bias difference computed over several hours, then it can be inferred that the inversion has incorrectly mapped the ionosphere. However, changes in the hardware of one of the receivers would also alter the receiver inter-frequency bias difference. This situation would be recognised as the difference should settle at its new value and remain constant. The receiver biases may also be affected by temperature, but this should be approximately common to site pairs.

Using the new method of evaluation, it has been shown that the full inversion technique offers improvements in the accuracy of TEC mapping compared to the thin shell. It is noteworthy that the strongest TEC gradients are found at higher latitudes, and the greatest improvements in the full inversion TEC mapping compared to the thin shell have been noted in this region. In considering these results, an important point to note is that it is solely the relative improvement in using the full inversion over the thin shell that is under consideration here. Absolute information about the accuracy of the ionospheric mapping cannot be inferred, only the TEC gradients between receiver pairs. Nevertheless, the work of this Chapter has indicated that, in the experimental case, the full four-dimensional inversion method offers advantages over the thin shell method, and has the potential to improve the accuracy of position solutions computed by single-frequency receivers.

6.6 Summary

In this Chapter, an experimental approach to the problem of determining the relative accuracy of the slant TEC measurement made by the thin shell and the full inversion methods has been described. The method, described by Meggs *et al.* (2004), is based on the expectation of the stability of the receiver inter-frequency biases. It relies on the assumption that the difference in the inter-frequency bias between two receivers is stable over a 24-hour period (Sardón and Zarraoa, 1997).

In the cases investigated, the inversion was always found to give an improvement over the thin shell for mid- to high-latitude receiver pairs. For low- to mid-latitude receiver pairs, the accuracy of the inversion was either comparable to, or better than, that attained using the thin shell method. Since the inversion method was shown generally to give smaller standard deviations than the thin shell, this suggests that, on an hour-by-hour basis, the full four-dimensional inversion can provide a more accurate representation of the ionosphere

Chapter 7

Comparisons with independent instrumentation

Abstract

Images of the electron density over northern Scandinavia in January 2002 are presented in this Chapter. The electron densities were produced using two independent techniques: MIDAS imaging and Incoherent Scatter Radar (ISR) observations. In the first case, observations of differential phase delays from GPS satellites were used to estimate the spatial distribution of electron density. In the second case, the EISCAT radar was used to gather independent data for comparison with the MIDAS images. The MIDAS results are presented in the form of latitude-altitude cross-sections of electron density at the locations and times coincident with the radar scans. The EISCAT data are plotted as "fan plots" showing electron density as a function of latitude and range. Also presented are wide-area maps of TEC for the entire day. The latitudinal positions and the time evolution of the trough seen in the MIDAS images are confirmed by the EISCAT radar. The study demonstrates the potential of MIDAS as a reliable tool for routine monitoring of the ionosphere on scale sizes comparable to those of the main trough.

7.1 Method

Since the early 1990s the EISCAT radar, illustrated in Figure 7.1, has provided independent measurements of electron density for the verification of tomographic images

Figure 7.1: The EISCAT radar UHF antenna at Tromsø.

constructed from carrier phase TEC measurements to NNSS satellites. The first independent EISCAT verification of a tomographic image was reported by Pryse and Kersley (1992). This work was followed in subsequent years with increasing sophistication, notably by Raymund *et al.* (1993) and Walker *et al.* (1996), who have demonstrated the value of independent measurements for the assessment of tomographic imaging under real ionospheric conditions. To date, MIDAS imaging using GPS TEC measurements has not been subjected to a comparison with incoherent scatter radar. A comparative study of this nature is described by Meggs *et al.* (2005, in press), and has led to greater confidence in the ability of MIDAS to reproduce large-scale structures in the ionosphere.

7.1.1 The EISCAT experiment

The EISCAT UHF radar at Tromsø (69.6°N, 19.2°E) was operated on 7 January 2002 in a special program mode, designated *tau3t_limb_UK@uhf*. This program, which is a modification of the UK EISCAT Special Program SP-UK-BLOB that was used by Pryse *et al.* (1996), is mainly used for scanning experiments where long range coverage is needed to get data from ionospheric altitudes at low-elevations. The radar experiment is based on a scan in the geographic meridian of Tromsø that starts at a low northward elevation (21° elevation) and scans southward in 55 steps of 0.152° latitude, which are defined at an altitude of 262.5 km. The scan dwells for 30 s at each step, and

the antenna requires a further 150 s to return to its starting position, making a total of 30 minutes to complete each scan. The *tau3t* alternating code transmitted during each north-to-south scan has a range resolution of 5.4 km, and is designed to provide information about the electron density, ion and electron temperatures, and line-of-sight plasma velocity over a range of 90 - 1400 km. A pre-integration period of 5 s is provided, and the data can be post-integrated to improve the data quality by adding two adjacent scanning positions. The radar data were analysed by colleagues from the EISCAT Group at the Rutherford - Appleton Laboratory (RAL).

The electron densities from the radar were calibrated using data from the Tromsø dynasonde. This was carried out by the EISCAT Group at RAL, Chilton. The dynasonde NmF2 are plotted against the EISCAT radar peak electron densities. The ionogram corresponding to any point in the dynasonde data is also displayed. Similarly, the points in the radar time series are accompanied by their vertical electron density profiles. This allows the operator to identify and reject any outlying points in the two time series, if necessary. A coefficient of calibration between the dynasonde and the radar data sets is then computed for the selected points. In the experiment under consideration here, the calibration coefficient was 1.466. The radar data are then re-analysed using this caalibration coefficient as a system constant. The fan plots shown in Figs. 7.3 to 7.6 were produced from this new data.

7.1.2 The MIDAS inversions

The MIDAS reconstructions were generated for the same period as the EISCAT experiment. By default, MIDAS maps from a voxel-based representation of the ionosphere to an ortho-normal representation using spherical harmonics in the horizontal domain, and ortho-normal basis functions in the vertical domain. The ortho-normal basis functions may be derived from models such as Chapman or Epstein. The inversions generated for this experiment were based on a vertical electron density profile formed using Chapman functions. The peak electron density, N_mF, and peak height, h_mF2, were estimated using the IRI-95 model at the horizontal centre of the grid. This gave values of 9.883×10^{11} electrons/m^3 and 284 km respectively. The electron density, N_H, at the scale height, H_mF2, was then found using the relation

$$N_H = N_mF2 \times \frac{1}{e} = 3.636 \quad [\times 10^{11} \text{ electrons m}^{-2}] \tag{7.1}$$

where $e = 2.718\ldots$.

The bottomside scale height H_mF2 was then found to be 68 km. The values of h_mF2 and H_mF2 were rounded to 280 km and 70 km respectively, and were set as the model

parameters that constrain the vertical electron density profile of the inversion. Four additional basis profiles were added: two at 250 km with scale heights of 60 km and 70 km; and two at 310 km with scale heights of 70 km and 80 km.

The GPS receiver data and satellite orbital data for the GPS receivers in Figure 7.2 were obtained from the SOPAC archive at http://sopac.ucsd.edu/ and used in the inversion. A global reconstruction grid of 1° in latitude and 4° in longitude, and 50 km in altitude, was set up in the MIDAS software. The imaging volume used extended from 30° N to 80° N, 10° W to 40° E, and from 80 km to 1180 km. The Kp indices for 7 January 2002, obtained from the World Data Centre (source: www.wdc.rl.ac.uk), ranged from 0 to 3, indicating quiet to moderate geomagnetic conditions. Under these conditions it can be expected that an empirical model such as IRI-95 would give a reasonable estimate of the peak and scale heights.

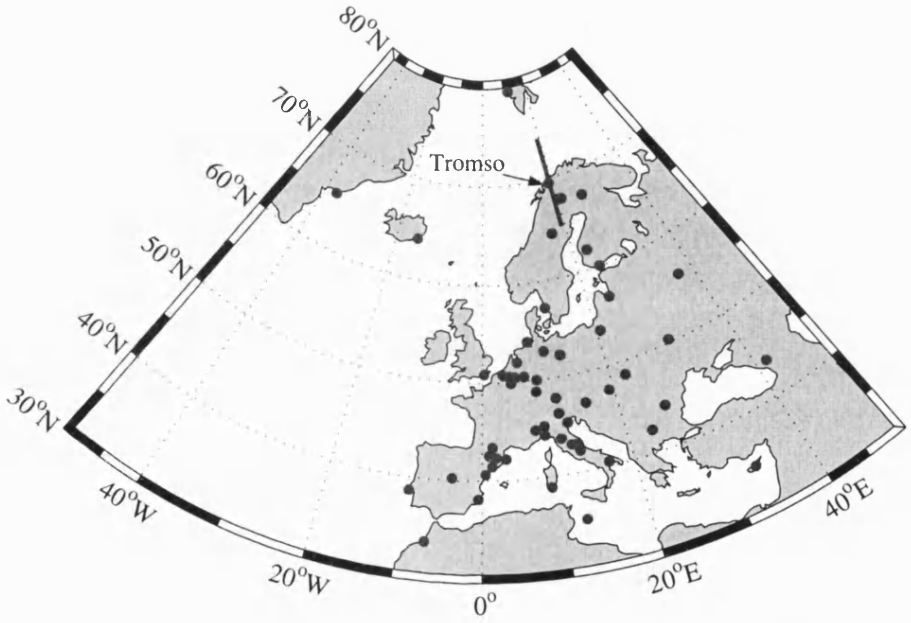


Figure 7.2: Map showing locations of GPS receivers and the EISCAT radar site at Tromsø. The line of the radar scan is shown passing through Tromsø.

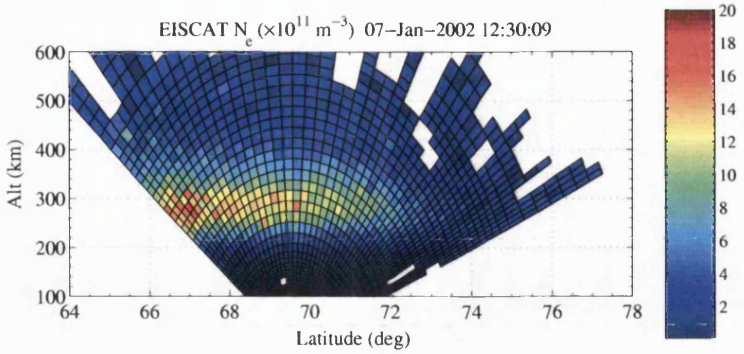
7.2 Results

MIDAS inversions result in one-hour movies of the three-dimensional distribution of electron density. These can be shown as vertical plots of electron density, or by integrating vertically through the inversion a map of vertical TEC can be produced. Here, single two-dimensional slices of the electron density at the longitude closest to the EIS-

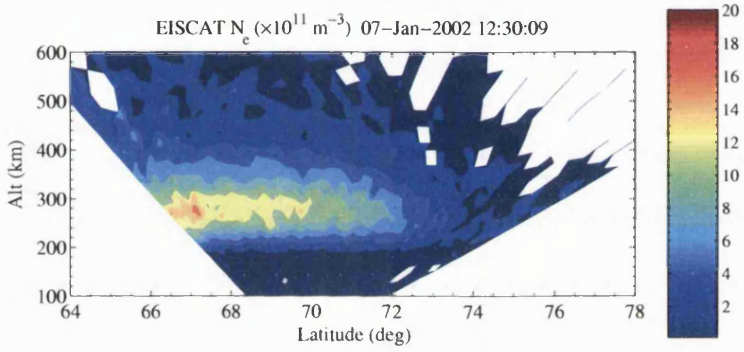
CAT radar are shown to permit visual comparison with the EISCAT images. Since the MIDAS imaging grid is stepped in increments of 4° longitude, the slice encompasses the longitude range from 18° E to 22° E. The latitudinal extent of the plots was limited to the region encompassing the EISCAT radar scans, that is from 64° N to 78° N. The times of the MIDAS images have been chosen to coincide with the half-way point of the EISCAT scan.

The electron density plots are presented in Figs. 7.3 to 7.6. There are three plots in each figure, showing (a) the uninterpolated EISCAT radar fan plots; (b) interpolated EISCAT radar fan plots; and (c) MIDAS electron density plots. The uninterpolated plots show all 50 radial scan lines with the radar returns scaled to electron density, mapped to axes of latitude and altitude. Thus the distribution of electron density can be readily seen. The interpolated plots show an average electron density distribution over the half-hour period of the scan, also mapped to axes of latitude and altitude. The MIDAS plots show electron density as a function of latitude and height at the time half way through the corresponding EISCAT scan.

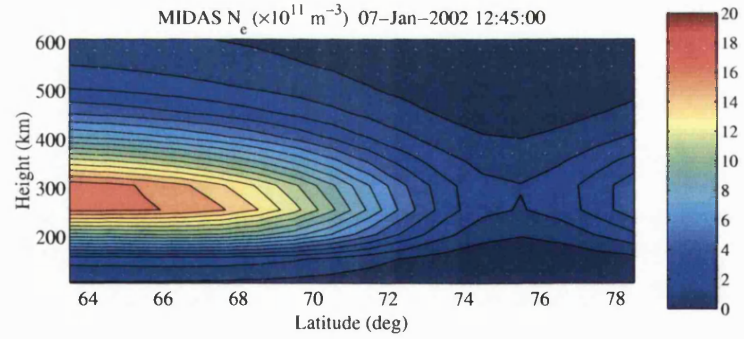
Figs. 7.3(a) and 7.3(b) show the electron density observed by the EISCAT radar for the half-hour starting at 12:30 UT, and Figure 7.3(c) shows the MIDAS electron density plot at 12:45 UT. The main dayside trough, centred around 75° N - 76° N, is clearly visible in all three images, with the equatorward wall at about 72° N. Figure 7.4 shows the situation between 13:00 and 13:30 UT. Both techniques reveal an overall decrease in the peak electron density in this half-hour period. Once again, there is close correspondence between the two techniques, although a temporary problem with the radar forced a premature halt to the scan. Following resumption of the scan, the half hour from 14:00 UT is shown in Figure 7.5, which displays a further decrease in the peak electron density. The two techniques both show a decrease in the electron density gradient at the equatorward wall compared to the earlier images. The main trough, centred around 74.5° N - 75.5° N, is again clearly visible in both techniques, and an apparent southward movement of the trough has also been reproduced. The southward movement of the trough is confirmed in Figure 7.6, in which it can be seen that both techniques reveal a further decrease in overall electron density. The equatorward wall of the trough is now at about 66° N. The EISCAT scans now show the trough minimum at about 74° N - about 1° further south than that shown in the MIDAS plots. The overall agreement between the two techniques in absolute values of electron density is of particular interest, and can be partly attributed to the correct calibration of the radar observations using the dynasonde. The vertical profile constraints that were computed from IRI-95 and used in MIDAS inversions appear to have been a suitable choice in this case, probably assisted by the quiet ionospheric conditions.



(a)

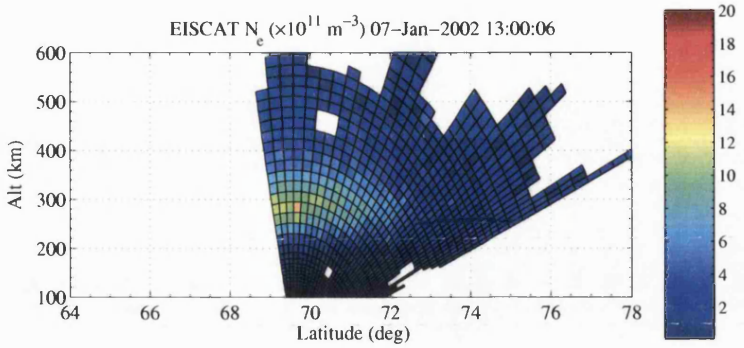


(b)

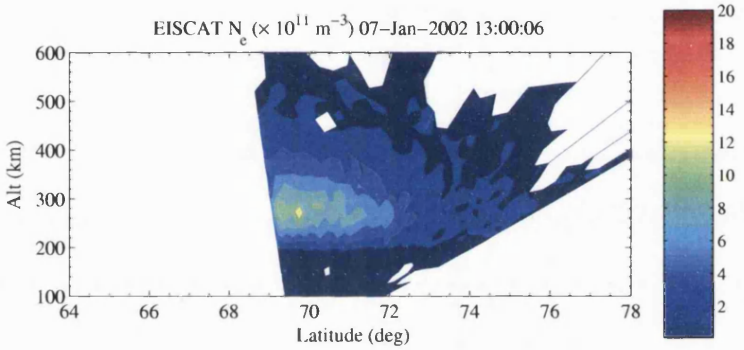


(c)

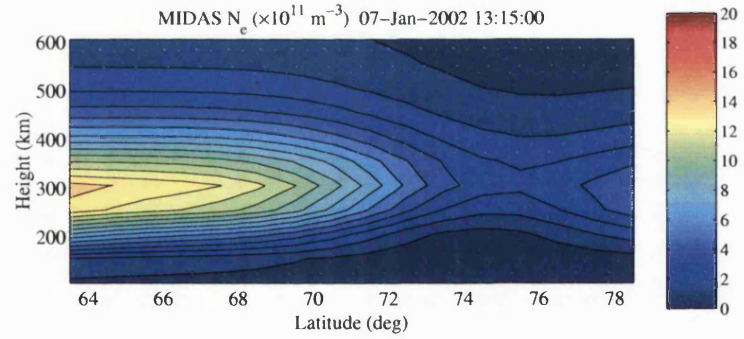
Figure 7.3: MIDAS Electron concentration at the longitude of the EISCAT radar between 12:30 and 13:00 UT: (a) Original EISCAT values, (b) Interpolated EISCAT values, and (c) MIDAS image. Figure constructed with the assistance of Dr. V. S. C. Howells, EISCAT Group, Rutherford-Appleton Laboratory.



(a)

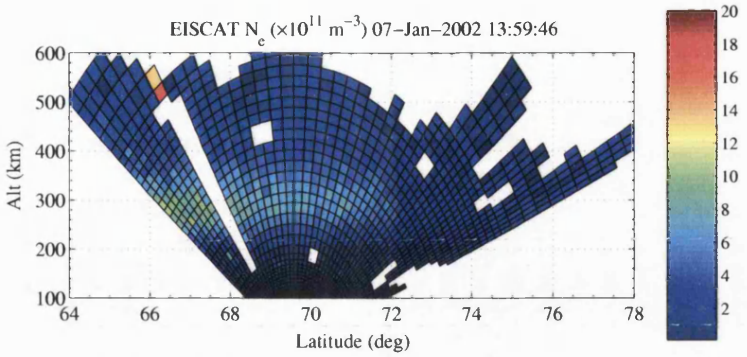


(b)

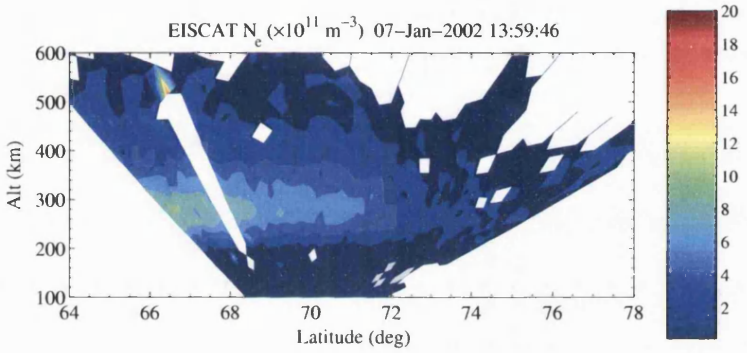


(c)

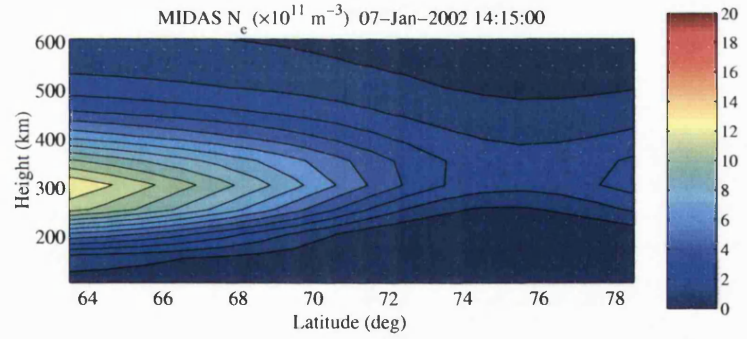
Figure 7.4: MIDAS Electron concentration at the longitude of the EISCAT radar between 13:00 and 13:30 UT: (a) Original EISCAT values, (b) Interpolated EISCAT values, and (c) MIDAS image. Figure constructed with the assistance of Dr. V. S. C. Howells, EISCAT Group, Rutherford-Appleton Laboratory.



(a)

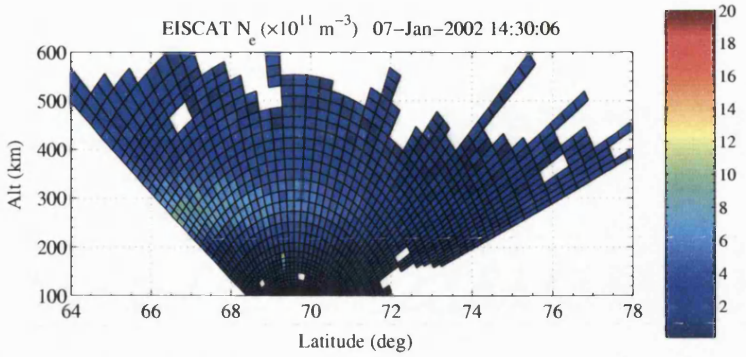


(b)

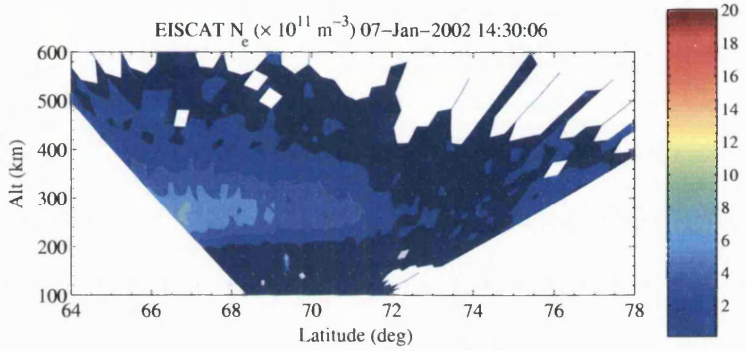


(c)

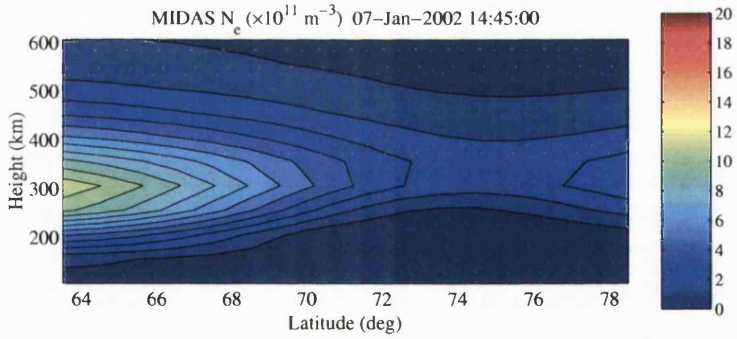
Figure 7.5: MIDAS Electron concentration at the longitude of the EISCAT radar between 14:00 and 14:30 UT: (a) Original EISCAT values, (b) Interpolated EISCAT values, and (c) MIDAS image. Figure constructed with the assistance of Dr. V. S. C. Howells, EISCAT Group, Rutherford-Appleton Laboratory.



(a)



(b)



(c)

Figure 7.6: MIDAS Electron concentration at the longitude of the EISCAT radar between 14:30 and 15:00 UT: (a) Original EISCAT values, (b) Interpolated EISCAT values, and (c) MIDAS image. Figure constructed with the assistance of Dr. V. S. C. Howells, EISCAT Group, Rutherford-Appleton Laboratory.

Figure 7.7 shows the vertical TEC at each hour from 12 to 17 UT on 7 Jan 2002. The diurnal variation of the TEC in this sequence can be seen to follow that expected from the solar radiation. The main trough is clearly visible as the longitudinally distributed depletion of electron density moving southwards from about 74° N at 12:00 UT to about 65° N at 17:00 UT.

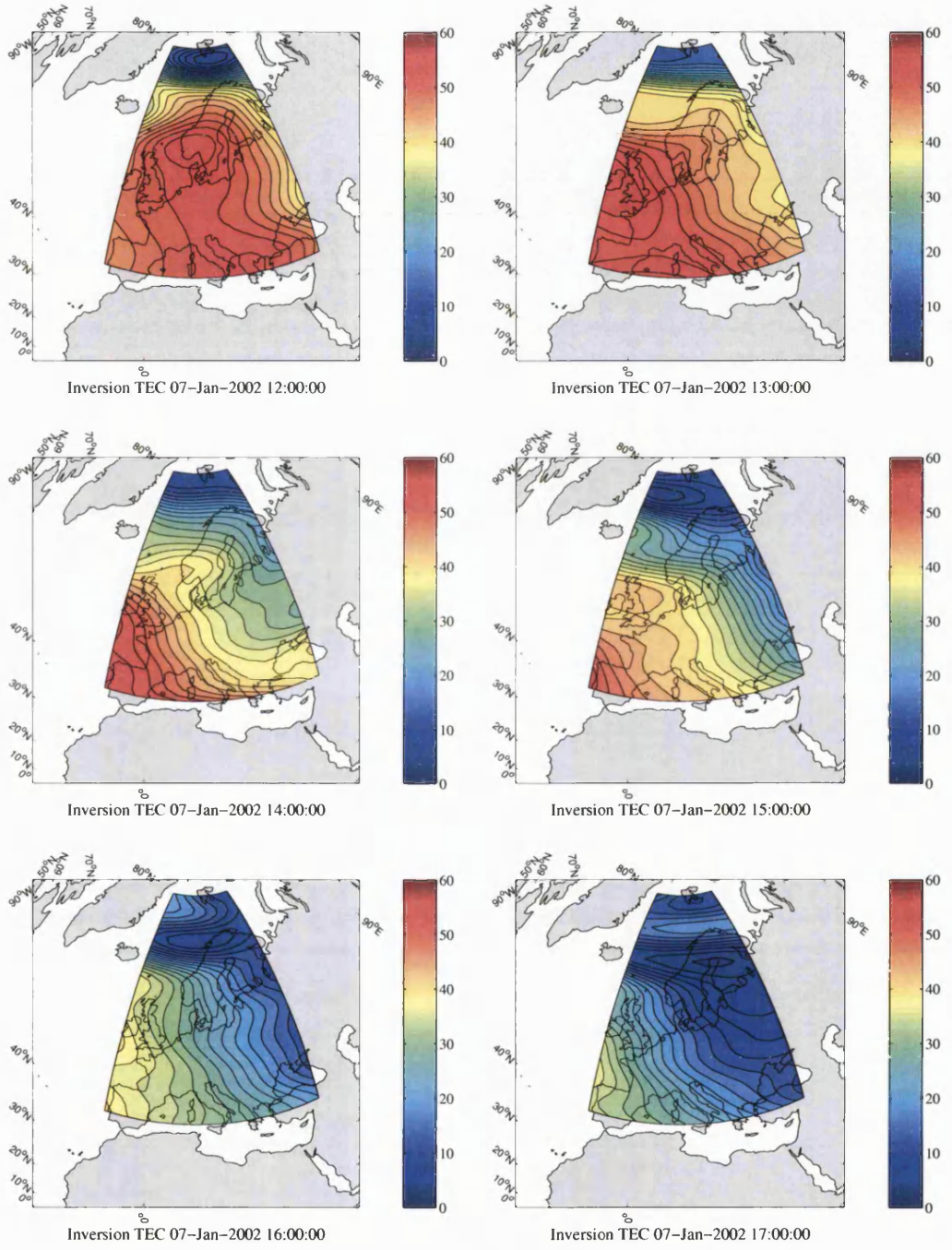


Figure 7.7: Post-noon changes in TEC showing the southward movement of the trough.

7.3 Discussion

In this chapter, images of the ionospheric electron density over Northern Scandinavia have been constructed using GPS imaging and incoherent scatter radar observations. The EISCAT radar has been extensively used to verify tomographic images of the ionosphere constructed from range measurements to NNSS satellites. (see, for example: Bernhardt *et al.* (1998), Foster, *et al.* (1994), Heaton *et al* (1995), Kersley *et al.* (1993), Mitchell *et al.* (1998), Pryse *et al.* (1996), Raymund *et al.* (1993) and Walker *et al.* (1996)). The work presented in this Chapter is believed to be the first independent verification of GPS imaging using the EISCAT radar.

Qualitatively, the results generally show good correspondence between the electron density images produced by the two techniques. However, a quantitative comparison is more difficult to achieve, since incoherent scatter radar and GPS imaging are different techniques. Moreover, the GPS images show constant electron density in voxels that are 4° longitude, 1° latitude and 50 km in height. This is well below the resolution available from incoherent scatter radar. In addition, the radar requires calibration external to the measurement itself (in this case a dynasonde was used).

The auroral ionosphere is a highly variable medium, and is subject to large changes in electron density on time scales less than that of the radar scan time (Kersley *et al.*, 1993). The EISCAT SP-UK-LIMB measurements are made over a 30-minute period, with each representing a spatially and temporally different 30 s snapshot of the ionosphere. Since the ionosphere cannot be assumed to be stationary over a 30-minute period, the fan plots represent the temporal and spatial evolution of the ionosphere over a 30-minute period. Conversely, the MIDAS electron density plots are instantaneous snapshots of the ionosphere.

In the full inversion, measurements of differential phase delay are inverted in a system of linear equations to solve for the underlying electron-density field. Since the equations are under-determined, model data is used to stabilise the solution. Models represent average or median conditions, and are less able to represent the highly variable conditions encountered at auroral latitudes. Nevertheless, during these geomagnetically quiet conditions the use of a model to stabilise the vertical profile in the imaging is valid because the vertical electron density profile at mid-latitudes (i.e., south of the trough) is well-modelled away from storm conditions. However, it is recognised that at later local times the auroral zone will extend further south and the imaging of the vertical profile may be less valid.

7.4 Summary

Two independent techniques have been used to image the winter time ionosphere over Northern Scandinavia. The first technique was MIDAS imaging using ground-based dual-frequency GPS carrier-phase data, and the second technique used incoherent scatter radar data from the EISCAT UHF radar at Tromsø, Norway. The vertical electron density profiles of the MIDAS inversions were constrained using peak and scale heights computed using the IRI-95 model. The EISCAT data were calibrated using data from the dynasonde located at the Tromsø site. Plots of electron density mapped to axes of latitude and altitude using both techniques have been presented, and demonstrate that the main trough can be reproduced accurately using data monitored by a network of ground based GPS receivers.

Chapter 8

The MIDAS Interface

Abstract

This chapter is concerned with the presentation of ionospheric information to users. The need for an intuitive user interface to MIDAS for a variety of different classes of users is discussed. The design and development of a basic graphical user interface (GUI) for use by non-expert operators will be described, and an example of a basic GUI will be presented and described. The basic GUI provides an indication of the vertical TEC at a specified location and time in terms of the statistics of the previous 30 days. Possible extensions to basic GUI that will provide an interface for expert operators to enter inversion parameters will also be discussed. During the development of the simple interface, a number of assumptions were made for programming convenience. These will be described in Subsection 8.2.1.

8.1 Design overview

Graphical computer interfaces, such as X-Window and Microsoft Windows, have evolved in response to the needs of non-specialist computer users. Such interfaces present users with an intuitive "front end" to applications, which then respond to events such as pointing and clicking with a mouse. Earlier interfaces were driven by a command line that required users to know a set of arcane commands and their options, which were entered verbosely using the keyboard. Matlab, the environment in which MIDAS is written, provides both a command line interface and the tools to produce bespoke GUIs. Commands can be run directly from the command line, or from within functions

and script files.

The MIDAS software was designed to take advantage of ionospheric observations from readily available instrumentation. It is a collection of Matlab programs that together implement a linear inversion algorithm to solve for the spatial and temporal distribution of electron density in the ionosphere. It has been extensively described by Mitchell and Spencer (2003), and a summary of the method has been given in Chapter 2. Input data from a variety of instruments are combined with inversion parameters to produce maps showing the horizontal distribution of vTEC, or vertical plots of electron density as a function of latitude and height. The resulting maps may be combined to produce movies showing the time evolution of the TEC or electron density structures.

In spite of the capabilities of Matlab to support a graphical style of interface, the MIDAS software currently retains the command line approach. The MIDAS algorithm is invoked by running the function invert.m in Matlab, using the following syntax:

$$\text{Result} = \text{invert}(\text{Input}, \text{Param})$$

It is possible to enter this directly on the Matlab command line, though it is more usual to run it within a script file that also sets the inversion parameters. The input and output arguments are Matlab structures, which can be stored in disc files. However, before the invert function can be called, the Input and Param structures must be generated. Input data can be derived from a variety of sources such as GPS (ground or space-based), ionosonde, NNSS or Topex, and the MIDAS suite includes functions and scripts that read the data files for each input type and form an Input structure. Where more than one data source is used, MIDAS contains a function to combine their data into a single Input structure suitable for use with the inversion function.

The Param structure contains all the inversion parameters. One of these is the geometry of the reconstruction grid, which is set using the function initgrid.m. Other parameters contained in this structure include the latitude and longitude limits, number of EOFs, model profiles, etc. The most efficient way to create the Param structure is to use a script file, which can also contain the call to invert.m. It is usual to limit the time duration of an inversion to one hour, so in order to invert longer periods of data it is necessary to batch inversions using suitable script files. This leads to a complex arrangement of inter-dependent script files and functions.

From the above description it can be appreciated that the effective use of the MIDAS software depends heavily upon the user's understanding of the underlying principles of the MIDAS algorithm and familiarity with Matlab. To change any inversion parameters or to add additional data sources it is necessary to edit Matlab scripts directly, thus

requiring some programming competence on the part of the user. With several different files to be edited "by hand", there is the possibility that the user will neglect to change an instance of a variable in a function, causing invalid calls or passing invalid data to other functions when the software is run. This will either lead to the abnormal termination of the software, or the software may appear to perform normally but will produce results that are meaningless. Worse, there may be no indication that the results are wrong.Two groups of users can be identified which, for the purposes of this work, can be classified as 'non-expert' and 'expert'. 'Non-expert' users are users who may be unfamiliar with programming in Matlab, and have little understanding of radio propagation and/or ionospheric physics. They may be specialists in other disciplines. A simple, intuitive interface which requires minimal input from the user and presents the results in a clear and easily interpreted manner is appropriate for this type of user. 'Expert' users can be considered to be users who are able to make informed decisions about input sources (GPS, ionosonde, etc) and inversion parameters, but may not necessarily be sufficiently familiar with programming languages to be confident about editing Matlab files. An advanced style of interface would be more appropriate for this class of user. This type of interface would enable expert users to set inversion parameters and to select input source data in a controlled manner. The two different styles of interface will be referred to in this work as 'basic' and 'advanced'. Whilst consideration has been given to both, actual development work was concentrated on the basic interface with the aim of producing a working prototype.

8.2 The 'Basic' Interface

8.2.1 The ionospheric activity indicator

In determining a suitable style of ionospheric activity indication for a non-specialist user, it was thought necessary to avoid directly giving the value of vertical TEC as a number. Instead, the indication is given as discrete levels ranging from 'very low' to 'very high', which are determined from the statistics of the preceding 30 days. Although the ionosphere is a highly variable medium, the short-term seasonal variability over periods of about 30 days can be considered linear for de-trending purposes. The approach adopted here takes as its input the date, time and position of interest and computes the vertical TEC. The corresponding vertical TEC values for the previous 30 days are obtained and put into a time series with the query day. A linear fit, y, is then computed:

$$\text{fit} = y = a_1 \times I_v(N) + a_0 \tag{8.1}$$

where I_v is the vertical TEC, N is the number of days preceding the query day ($N = 0$ being the query day), and a_0 and a_1 are the coefficients of the linear fit. An example showing the vertical TEC values and the fit line for 12:00 UT on 1st May 2002 at Chilton (UK) are plotted in Fig. 8.1(a).

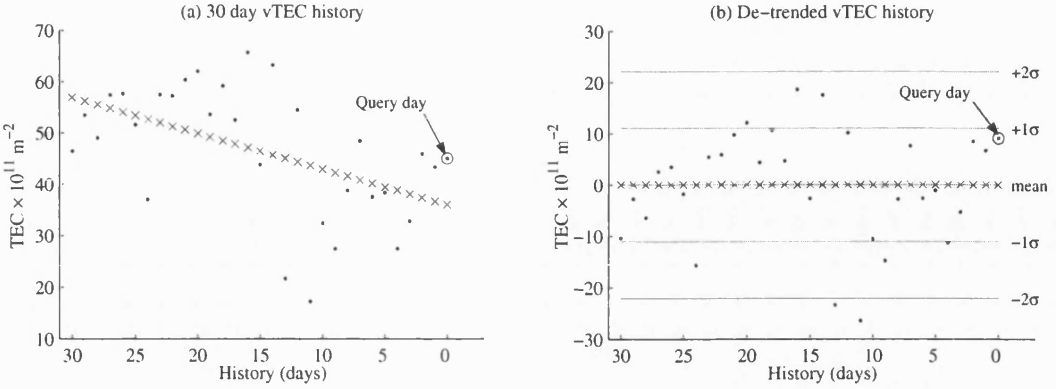


Figure 8.1: (a) Vertical TEC values for the query day (1 May 2002) and the preceding 30 days (dot) and the linear fit line (cross), and (b) the data after de-trending. The mean and standard deviation (σ) lines are indicated.

The TEC data were then de-trended by computing the differences between the TEC values and their corresponding points on the fit line (see Fig. 8.1(b)). The thresholds between the different levels of TEC were then computed using the standard deviation of the differences for the preceding 30 days, obtained from:

$$\sigma = \mathrm{std}\left[\mathrm{I_v}\left(N\right) - y\left(N\right)\right] \quad \text{with} \quad N > 0 \tag{8.2}$$

This number was then compared with a test number, q, obtained from the vertical TEC on the query day using

$$\mathrm{q} = \frac{\mathrm{I_v}\left(0\right) - y\left(0\right)}{\sigma} \tag{8.3}$$

and the level of ionisation in the ionosphere was determined using the following scale of thresholds, which are shown in Fig. 8.1(b):

q less than -3σ	very low TEC
q between -3σ and -2σ	low TEC
q between -2σ and -1σ	medium low TEC
q between -1σ and $+1\sigma$	medium TEC
q between $+1\sigma$ and $+2\sigma$	medium high TEC
q between $+2\sigma$ and $+3\sigma$	high TEC
q more than $+3\sigma$	very high TEC

8.2.2 Description of 'Basic' Interface

The 'basic' interface was written using the Matlab GUI designer GUIDE, and is shown in Fig. 8.2. In use, the user specifies a date and hour using the four popup menus at the upper left, then presses the 'Get Data File' button. The program then searches for the required result data and reports on its availability in the 'messages' box.

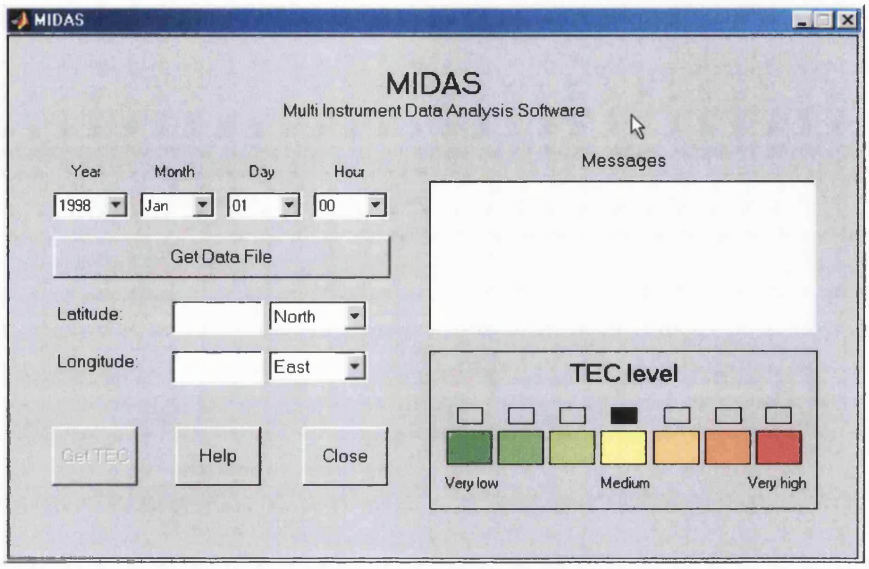


Figure 8.2: The 'Basic' MIDAS interface.

If the result file is available, the 'Get TEC' button becomes enabled. The user now enters a geographical position in the 'Latitude' and 'Longitude' edit boxes. Pressing the 'Get TEC' button will load the result data for the query day and for the previous 30 days and compute the vertical TEC values. Using the procedure described in the previous section, the level of ionospheric activity is indicated on the scale at the lower right by the position of the black square. An additional feature of the interface allows the user to obtain the maximum value of vertical TEC, and report it and its geographical position in latitude and longitude by leaving both the 'Latitude' and 'Longitude' edit boxes blank and pressing the 'Get TEC' button after loading the result data.

A short help message is available by clicking the 'Help' button. This appears as a separate window with an 'OK' button to allow the user to dismiss it, although it is not modal and hence the help window can remain visible whist the interface is being used. In the event of either of the 'Latitude' or 'Longitude' boxes being left blank, a modal message window will appear to warn the user that one of these values is missing. The user cannot continue until this message has been dismissed by clicking on its 'OK' button. The interface can be closed by clicking its 'Close' button and answering 'Yes' to the confirm dialog that appears. Most controls in the interface have associated with

them a textual hint on their use, called a 'tooltip', that appears when the mouse pointer is allowed to hover above the control.

8.3 Further development

The method described in Section 8.2.1 represents an approach to the problem of determining whether the level of ionisation in the near-space plasma lies within acceptable limits for the user. The first stage in refining this method would be to test the validity of the assumptions made in Section 8.2.1 and, if necessary, change the interface code to accommodate this. For example, a "caution" level could be added by setting an intermediate threshold. Further discussions with BAE Systems will be necessary to determine any other requirements of the user interface, although there appear to be security constraints that would limit the amount of detail that they could provide.

Default values and parameters could be set in a configuration file that would be accessible only to an expert operator. Access control will be easily achieved in a Unix environment using the standard Unix file attributes. Other external files could include a "last state" file. On closure of the interface, the last state, i.e. the date, time and position would be saved to the "last state" file. On the next startup, these values would be loaded into the interface and would be immediately accessible to the user.

In future, an advanced MIDAS interface, aimed at the fully expert user, would be developed that would permit comprehensive setting of all reconstruction parameters and data sources. This would, however, require extensive consultation with the user community and is beyond the scope of this project.

8.4 Summary

In this chapter, a solution to the problem of presenting inversion data to users has been considered. Graphical User Interfaces (GUIs) in general were discussed, and the reasons for providing one for the MIDAS program were presented. In its current form, MIDAS retains a command-line style of interface that requires users to know a set of arcane commands that are entered using the keyboard. To use MIDAS effectively requires an understanding of the inversion algorithm and familiarity with Matlab programming. There exists a danger that incorrect entry of inversion parameters would cause the software either to terminate abnormally or to produce meaningless results, and for the user to be unaware that the results are meaningless. Thus two different classes of

user were identified, who were referred to as 'expert' and 'non-expert'. The interface described in this Chapter is designed to provide non-expert users with an intuitive way of interacting with the program and an easily understood indication of the level of ionisation in the ionosphere.

A working prototype interface, termed 'Basic' and intended for 'non-expert' users, was produced. Instead of giving a numerical value of TEC, it uses a moving block to indicate seven discrete levels of TEC ranging from 'very low' to 'very high'. The levels are determined from the statistics of the preceding 30 days, and assume that the short-term seasonal variability of the ionosphere is linear over this period. An example of a basic style of interface was then presented, and an explanation of its use was given. The chapter concludes with a short discussion on the possible future development of the MIDAS interface.

Chapter 9

Conclusions and Future Work

9.1 Conclusions

When this research project was commenced, a new technique for imaging the electron density in the ionosphere had already been introduced by Spencer and Mitchell (2001). Using GPS dual-frequency data, which has been readily available since about 1995, the new technique can be used to image large areas of the ionosphere at low resolution. It can give an overview of large-scale gradients and convection cells. The global extents and time evolution of ionospheric storms can also be studied using the technique. Although this new technique had been developed and tested on experimental GPS data, it was largely untested and had not been focussed onto any specific applications. The purpose of the work here was to:

1. verify the technique by simulation.
2. quantify the advantages of the new technique over existing methods
3. verify the use of the technique in an experimental case.

The imaging was limited to the European region for a number of reasons. Firstly, ground-based GPS instrumentation is distributed over the land masses, and the continental European region is particularly well provided with suitable instrumentation. Secondly, a requirement of the industrial sponsor was for ionospheric specification over the European region. Thirdly, the EISCAT radar is situated in northern Europe, and this instrument was ideal for providing independent verification in the experimental environment.

An initial feasibility study was presented, in which the new full inversion technique was compared to the established thin shell method. The results of applying the two techniques were quantified by determining the improvements in ionospheric delay correction for two fictitious single-frequency GPS users, located on mainland Europe in Hamburg and Milan. It was found that the new technique yielded line of sight ionospheric delays with errors consistently less than those found using the established thin shell method.

A more extensive simulation study was presented, which was designed to verify the accuracy of the new technique. The ionosphere in the simulation study was based on the IRI-95 model, and actual satellite and receiver geometries were used. The study showed that vertical TEC errors resulting from the full MIDAS inversion were less than 15%. In contrast, the largest errors resulting from the established thin shell method could be as high as 90%. Generally, a threefold improvement was found when using the full inversion technique, in comparison to the thin shell. The simulation study also revealed that the full inversion was able to reproduce small-scale horizontal ionospheric structures, whereas the thin shell tended to smooth them out.

When using real data, it is more difficult to quantify the errors since there is no truth ionosphere with which to make comparisons. To deal with this, two different approaches to quantify the slant TEC error were adopted. The first was to use the differential code method of determining slant TEC along each individual satellite-to-receiver path and to correct these for the inter-frequency biases. These estimates of slant TEC were then taken as true measurements the ionosphere, and were compared to the corresponding integrations through the MIDAS images in order to evaluate the image accuracy. This approach yielded estimated slant TEC errors through the MIDAS images in the order of 0.3 TECu.

The second approach to evaluating the new technique made use of a known quality of the inter-frequency biases, namely their stability. Each individual MIDAS inversion used one hour of data, and the residuals between the initial differential code observations and the corresponding paths through frames of the MIDAS image is an instantaneous measure of the combined satellite-to-receiver inter-frequency bias. If this measure of the inter-frequency bias were to remain constant hour by hour, then the implication is that the MIDAS images of TEC are correct. In other words, the error in TEC maps directly into the error in the inter-frequency bias. This measure was exploited to make a comparison experimentally between the established thin shell method and the new MIDAS inversion technique. It was found that in all cases the full inversion was either equivalent to, or better than, the thin shell by this metric. The greatest improvements were found at higher latitudes where stronger horizontal gradients in TEC are found. This was consistent with the results of the simulation study, wherein the thin shell had been shown artificially to smooth the horizontal structures.

The verification of the technique in terms of the spatial distribution of the electron density was accomplished using the EISCAT radar. During January 2002 a special meridional scanning experiment was carried out using the radar, and observed the main trough. This feature was seen to migrate south over successive scans. The MIDAS inversion was able to provide images which compared well with the EISCAT results. The main trough was clearly visible in each of the MIDAS results. Much smaller features were visible in the radar data, but it was not possible to verify whether these were below the resolution of the MIDAS technique. For the present experimental configuration, the MIDAS technique is not limited in resolution by the inversion itself, but rather by the density of observation data in Northern Scandinavia. In principle, the MIDAS inversion is capable of imaging structures down to those scales previously reported by other researchers. See, for example, Mitchell *et al.* (1997) and Biswas and Na (2000).

At this stage the MIDAS software was still under development and consequently was not suitable for routine operational use. However, since the method had been tested and developed to a point where it could be automated, the technique could be used for routine ionospheric monitoring by a non-expert user through a graphical user interface. Discussions with the industrial sponsor guided the requirements for the prototype interface, which is shown in Figure 8.2. The requirements agreed upon were, simply, to classify the TEC at a given location in Europe at a specified time on a particular day. The classification used consisted of 7 bands relating to the standard deviation of the TEC in comparison to the de-trended TEC for the previous 30 days at that particular time of day. This would tell the user whether the TEC was anomalously low or high, and could be used as a diagnostic for possible problems with HF instrumentation.

9.2 Future work

Plans for future work involve the development of an effective method of ionospheric forecasting using the MIDAS images. The most challenging part is to predict ionospheric storm activity.

The MIDAS algorithm is already being applied in pure scientific studies of the ionosphere. In particular, GPS imaging is the only method whereby large regions of the ionosphere can be viewed continuously to study ionospheric storms. Other work focuses on the relationships between the large-scale TEC and small-scale ionospheric irregularities causing both phase and amplitude scintillation.

This research project has shown that the thin shell method of TEC mapping does not

cope well with strong horizontal TEC gradients. WAAS and similar aircraft navigation projects make use of the thin shell method. A particular problem confronting the development of a wide-area ionospheric correction system in the vicinity of the equatorial anomaly (that is, over South America) is that the horizontal gradients can be very strong. Similarly, in high latitude regions large TEC gradients and high variability can be encountered. Work is planned that will seek to apply the new imaging technique in regions such as these.

In summary, this research has brought the technique forward from being a largely untested method to the point at which it could be used and relied upon operationally for non safety critical applications.

References

Austen J. R., Franke S. J., Liu C. H. and Yeh K. C. "Application of computerized tomography techniques to ionospheric research", *Proc. Beacon Satellite Symposium 1986 (Ed: A. Tauriainen)*, ISBN 951-42-2256-3, University of Oulu, Finland, pp 25-35, 1986.

Austen J. R., Franke S. J. and Liu C. H. "Ionospheric imaging using computerized tomography", *Radio Science*, Vol. 23, No. 3, pp 299-307, May/June 1988.

Beynon W. J. G. and Williams P. J. S. "Inchoherent scatter of radio waves from the ionosphere", *Rep. Prog. Phys.*, 41, pp 910-956, 1978.

Bernhardt P. A., McCoy R. P., Dymond K. F., Picone J. M., Meier R. R., Kamalabadi F., Cotton D. M., Charkrabarti S., Cook T. A., Vickers J. S., Stephan A. W., Kersley L., Pryse S. E., Walker I. K., Mitchell C. N., Strauss P. R., Na H., Biswas C., Bust G. S., Kronschnabl G. R. and Raymund T. D. "Two-dimensional mapping of the plasma density in the upper atmosphere with computerized ionospheric tomography (CIT)", *Physics of Plasmas*, 5, 5, pp. 2010 - 2021. 1998.

Bilitza D.,"The International Reference Ionosphere 1990," Natl. Space Sci. Data Cent./World Data Cent. for Rockets and Satell. 90-22, Greenbelt, Md.

Birch M. J., Hargreaves J. K. and Bailey G. J., "On the use of an effective ionospheric height in electron content measurement by GPS reception", *Radio Science*, Vol. 37, No. 1, 10.1029/2000RS002601, 2002.

Biswas C. and Na H., "Resolution and coverage analysis for ionospheric tomography", *Radio Science*, 35(3), pp 905 - 920, 2000.

Ciraolo L. and Spalla P. "Comparison of ionospheric total electron content from the Navy Navigation Satellite System and the GPS", *Radio Science*, Vol. 32, No. 3, pp 1071-1080, May-June 1997.

Ciraolo L. and Spalla P. "Comparison of TEC from NNSS and GPS from 1994 to 2000", *Proc. of International Beacon Satellite Symposium*, Boston, USA, 4-6 June 2001.

Coco D. S., Coker C., Dahlke S. R. and Clynch J. R., "Variability of GPS satellite differential group delay biases", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 27, No. 6, November 1991.

Collis P. N. and Haggstrom I., "Plasma convection and auroral precipitation processes associated with the main ionsopheric trough at high latitudes" *J. Atmos. Terr. Phys.*, Vol. 50(4/5), pp 389-404.

Davies, K. "Ionospheric Radio", Peter Peregrinus Ltd., London, 1990. ISBN 086341186X.

Davies K. and Hartmann G. K. "Studying the ionosphere with the Global Positioning System", *Radio Science*, Volume 32, Number 4, pp 1695-1703, July/August 1997.

Duncan C., Spitzmesser D., Meehan T., Franklin G. and Wilson B "An interfrequency bias calibrator for TurboRogue GPS receivers", *Proc. of Inst. of Navigation National Technical Meeting* 1998, pp 659 - 666.

El-Arini, B. M., Poor W., Lejeune R., Conker R., Fernow J. and Markin K. "An introduction to Wida Area Augmentation System and its predicted performance", *Radio Science*, Volume 36, Number 5, pp 1233-1240, September/October 2001.

Foster J.C., Buonsanto M.J., Holt J.M., Klobuchar J.A., Fougere P., Pakula W., Raymund T.D., Kunitsyn V.E., Andreeva E.S., Tereshchenko E.D. and Khudukon B.Z. "The Russian-American tomography experiment", *Int. J. of Im. Syst. and Tech.*, Vol. 5, Iss. 2, pp 148 - 159, 1994.

Fougere P.F. "Ionospheric radio tomography using maximum entropy 1. Theory and simulation studies", *Radio Science*, Volume 30, Number 2, pp 429-444, March/April 1995.

Fremouw E. J., Secan J. A. and Howe B. M. "Application of stochastic inverse theory to ionospheric tomography", *Radio Science,* Volume 27, Number 5, pp 721-732, September/October 1992.

Gao Y. and Liu Z. Z., "Precise ionosphere modelling using regional GPS network data", *Journal of Global Positioning Systems*, Vol. 1, No. 1, pp 18-24, 2002.

Hargreaves J. K. "The solar-terrestrial environment", ISBN 0-521-42737-1, Cambridge University Press, 1992.

Harris I. L., Mannucci A. J., Iijima B. A., Lindqwister U. J., Muna D., Pi X. and Wilson B. D, "Ionospheric specification algorithms for precise aircraft GPS-based aircraft navigation", *Radio Science*, Vol. 36, No. 2, pp 287-298, March/April 2001.

Heaton J.A.T., Pryse S.E. and Kersley L., "Improved background representation, ionosonde input and independent verification in experimental ionospheric tomography", *Annales Geophysicae*, Vol. 13, Iss. 12, pp 1297 - 1302, 1995.

Hofmann-Wellenhof B., Lichtenegger H. and Collins J, "GPS Theory and Practice" 5th edition, ISBN 3-211-83534-2, Springer-Verlag, Wien New York, 2001.

Kaplan E. D., Leva J. L. and Pavloff M. S. In: "Understanding GPS: Principles and Applications" (Elliott D. Kaplan, ed.) ISBN 0890067937, Artech House, Boston, MA. 1996.

Kee, Changdon and Parkinson, Bradford W. "Wide Area Differential GPS (WADGPS): Future Navigation System", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 32, No. 2, April 1996.

Kersley L., Heaton J. A. T., Pryse S. E. and Raymund T. D., "Experimental ionospheric tomography with ionosonde input and EISCAT verification", *Ann. Geophysicae* Vol. 11, pp. 1064 - 1074, 1993.

Klobuchar, John A., "Ionospheric time-delay algorithm for single-frequency GPS users", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. **AES-23**, No. 3, May 1987.

Lanyi G. E. and Roth T. "A comparison of mapped and measured total ionospheric electron content using global positioning system and beacon satellite measurements", *Radio Science*, Vol. 23, No. 4, pp 483-492, July/August 1988.

Mannucci A. J., Wilson B. D. and Edwards C. D. "A new method for monitoring the Earth's ionospheric total electron content using the GPS global network", *Proc. of the Institute of Navigation GPS Meeting* 1993, pp. 1113-1122.

Mannucci A. J., Wilson B. D., Yuan D. N., Ho C. H., Lindqwister U. J. and Runge T. F., "A global maping technique for GPS-derived ionospheric total electron content measurements", *Radio Science*, Volume 33, Number 3, pp 565-582, May-June 1998.

Mannucci A. J., Iijima B. A., Lindqwister U. J., Pi X., Sparks L. and Wilson B. D., "GPS and Ionosphere" in: Review of Radio Science 1996 - 1999 (W. Ross Stone, ed.), Oxford University Press, 1999. ISBN 0198565712.

McCoy R.P., "Using space-based remote sensing for improved global navigation and communication", *IEEE International Geoscience And Remote Sensing Symposium, Vols I - VII, Proceedings - Learning from Earth's Shapes and Sizes*, 1016-1018, 2003

McNamara Leo F. "The Ionosphere: Communications, Surveillance and Direction Finding", Krieger, Malabar, Florida, USA, 1991. ISBN 0894640402.

Meggs R. W., Mitchell C. N. and Spencer, P. S. J. "A comparison of techniques for mapping total electron content over Europe using GPS signals", *Radio Science*, 39, RS1S10, doi:10.1029/2002RS002846, Feb. 2004.

Meggs R. W., Mitchell C. N. and Howells V. S. C.. "Simultaneous observations of the main trough using GPS imaging and the EISCAT radar", *Annales Geophysicae*, in press.

Mitchell C. N., Kersley L. and Pryse S. E. "The effects of receiver location in two-station experimental ionospheric tomography", *J. of Atmos. and Sol. Terr. Phys.*, Vol. 59, No. 12, pp 1411 - 1415, 1997.

Mitchell C. N., Walker I. K., Pryse S. E., Kersley L., McCrea I. W. and Jones T. B. "First complementary observations by ionospheric tomography, the EISCAT Svalbard radar and the CUTLASS HF radar", *Ann. Geophysicae* 16, pp1519-1522, 1998.

Mitchell C. N. "Imaging the near-Earth space plasma", *Phil. Trans. R. Soc. Lond.* A (2002) 360, 1 -14.

Mitchell C. N. and Spencer, P. S. J. "Ionospheric electron concentration mapping using GPS over Europe and USA during the storm of July 2000", in *Proc. of the 27th General Assembly of teh International Union of Radio Science, Paper 1375* [CD-ROM].

Mitchell C. N. and Spencer, P. S. J. "A three-dimensional time-dependent algorithm for ionospheric imaging using GPS", *Annals of Geophysics*, 46(4), pp687-696, August 2003.

Na H. and Lee H., "Resolution analysis of tomographic reconstruction of electron density profiles in the ionosphere", *Int. J. of Im. Syst. Technol.*, Vol. 2, pp 209-218, 1991.

Na H. and Sutton E., "Resolution analysis of ionospheric tomography systems", *Int. J. of Im. Syst. Technol.*, Vol. 5, pp 169-173, 1994.

Na H., Hall B. and Sutton E., "Ground station spacing effects in ionospheric tomography", *Ann. Geophysicae* Vol. 13, pp. 1288 - 1296, 1995.

Parkinson B.W., Spilker J.J., Axelrad P. and Enge, P. (Eds) "Global Positioning System: Theory and applications, volume 1", ISBN 1-56347-106-X, American Institute of Aeronautics and Astronautics, 1996.

Pryse S.E. and Kersley L,"A preliminary experimental test of ionospheric tomography", *J. Atmos. Terr. Phys.* Vol. 54, No. 7/8, pp 1007 - 1012, 1992.

Pryse S.E., Kersley L. and Walker I.K. "Blobs and irregularities in the auroral ionosphere", *J. of Atmos. and Terr. Phys.*, Vol.58, No. 1 - 4, pp. 205 - 215, 1996.

Pryse S.E., Kersley L., Mitchell C.N., Spencer P.S.J. and Williams M.J. "A comparison of reconstruction techniques used in ionospheric tomography", *Radio Science*, Vol. 33, No. 6, pp 1767 - 1779, Nov/Dec 1998.

Raymund T. D., Austen J.R., Franke S. J., Liu C. H., Klobuchar J. A. and Stalker J. "Application of computerized tomography to the investigation of ionospheric structures", *Radio Science*, Vol. 25, No. 5, pp 771 - 789, Sept/Oct 1990.

Raymund T. D., Pryse S. E., Kersley L. and Heaton J. A. T. "Tomographic reconstruction of ionospheric electron density with European incoherent scatter radar verification", *Radio Science* 28, No. 5, pp 811-817, Sept/Oct 1993.

Raymund T. D. "Comparisons of several ionospheric tomography algorithms", *Annales Geophysicae*, **13**, pp 1253 - 1262, 1995.

Reinisch B. W., Huang X., Belehaki, A. and Jodogne, J-C., "Bottom and topside ionospheric TEC obtained from ground-based ionosonde measurements", *Proc. of Int. Beacon Satellite Symposium*, Boston, USA, 4-6 June 2001.

Rishbeth H. In: "Propagation of Radio Waves", (Hall, M. P. M., Barclay, L. W. and Hewitt, M. T., eds.) ISBN 0-85296-819-1. The Institution of Electrical Engineers, Stevenage, Herts, UK. 1996

Rishbeth H. and Williams P. J. S. "The EISCAT ionospheric radar - the system and its early results", *Quart. J. Roy. Astron. Soc.*, Vol. 26, Iss 4, pp478-512, 1985.

Sardón E. and Zarraoa N., "Estimation of total electron content using GPS data: how stable are the differential satellite and receiver instrumental biases?", *Radio Science*, Vol. 32, No. 5, pp 1899-1910, September/October 1997.

Schreiner W. S., Markin R. E. and Born G. H., "Correction of single frequency altimeter measurements for ionosphere delay", *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 35, No. 2, March 1997.

Spencer P. S. J. and Mitchell C. N., "Multi-instrument inversion technique forionospheric imaging", *Proc. of the International Beacon Satellite Symposium*, Boston, 2001, pp 25 -35.

Titheridge J. E., "Determination of ionospheric electron content from the Faraday rotation of geostationary satellite signals", *Radio Science*, Vol. 20, pp 353-369, 1972.

Walker I. K., Heaton J. K., Kersley L., Mitchell C. N., Pryse S. E., and Williams m. J. "EISCAT verification in the development of ionospheric tomography", *Ann. Geophysicae* 14, 1413-1421 (1996).

Wilson B. D., Mannucci A. J. and Edwards C. D. "Sub-daily northern hemisphere ionospheric maps using the IGS GPS network", *Proc. of the 7th International Ionospheric Effects Symposium*, J. Goodman, ed., Alexandria, VA., USA, May 1993.

Wilson B. D., Mannucci A. J. and Edwards C. D. "Sub-daily northern hemisphere ionospheric maps using an extensive network of GPS receivers", *Radio Science*, Vol. 30, No. 3, pp 639-648, May/June 1995.

Woo K. T. "Optimum Semicodeless Carrier-Phase Tracking of L2", *Navigation: Journal of the Institute of Navigation*, Vol. 47, No. 2, Summer 2000.

Yeh K. C. and Raymund T. D. "Limitations of ionospheric imaging by tomography", *Radio Science*, Volume 26, Number 6, pp 1362-1380, November/December 1991.

Yin P., Mitchell C.N., Spencer P.S.J. and Foster J.C. "Ionospheric electron concentration imaging using GPS over the USA during the storm of July 2000", *Geophysical Research Letters* Vol. 31, L12806, doi:10.1029/2004GL019899, 2004.

Yunck, Thomas P., Liu, Chao-Han and Ware, Randolph, "A History of GPS Sounding", *Terr. Atmos. Ocean Science*, Volume 11, No. 1, pp. 1 - 20, 2000.

Appendices

Appendix A

Maps of vertical TEC

In this appendix, a complete set of vertical TEC maps for each hour of the 25-day data set is presented. These TEC maps have been constructed from inversions of real sTEC measurements through the ionosphere. The same colour range is used for all maps to facilitate visual comparison. The images progress in time from left to right, starting with 00:00 UT at top left and concluding with 23:00 UT at bottom right.

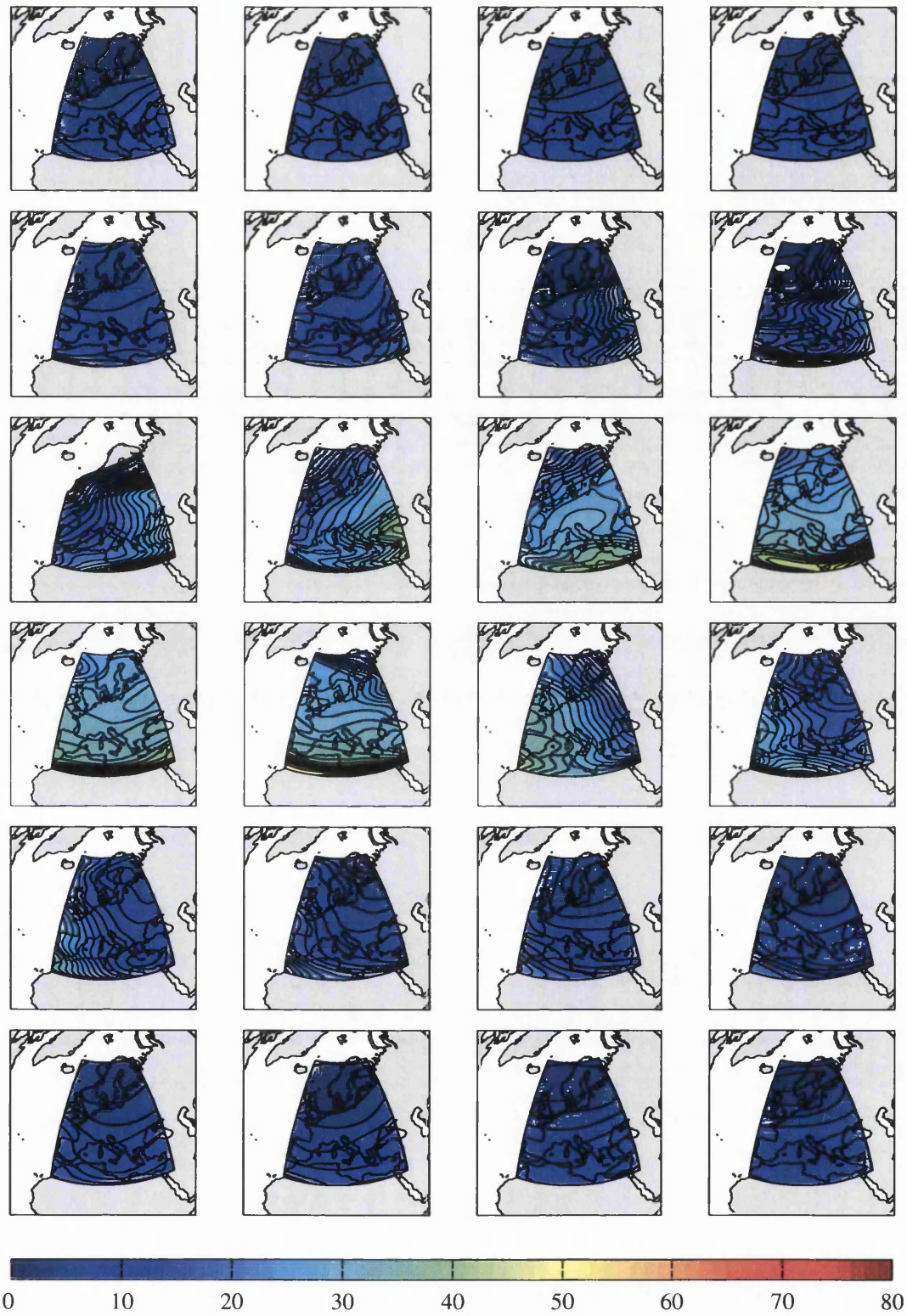


Figure A.1: Hourly vertical TEC maps for 01 January 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

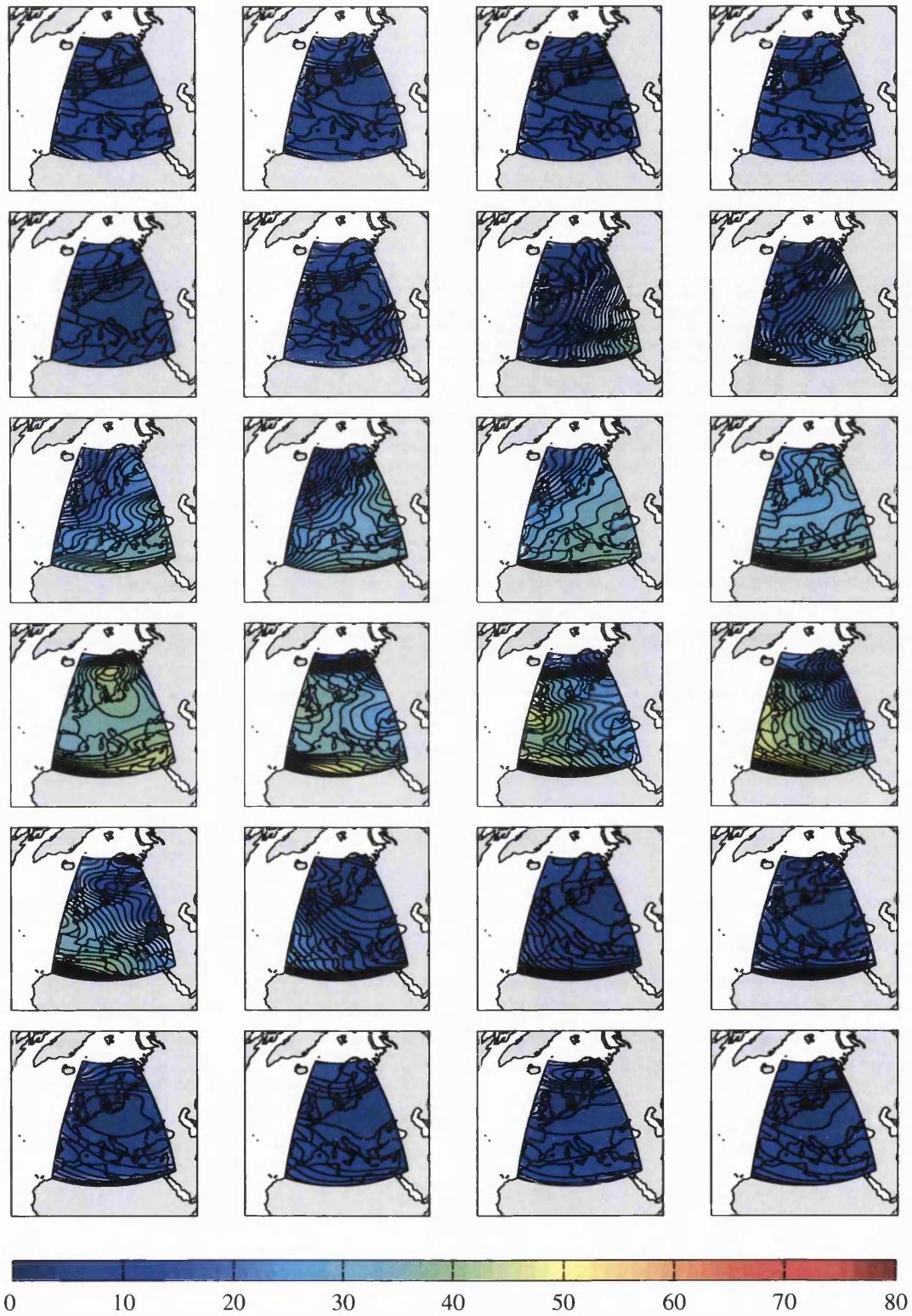


Figure A.2: Hourly vertical TEC maps for 24 January 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

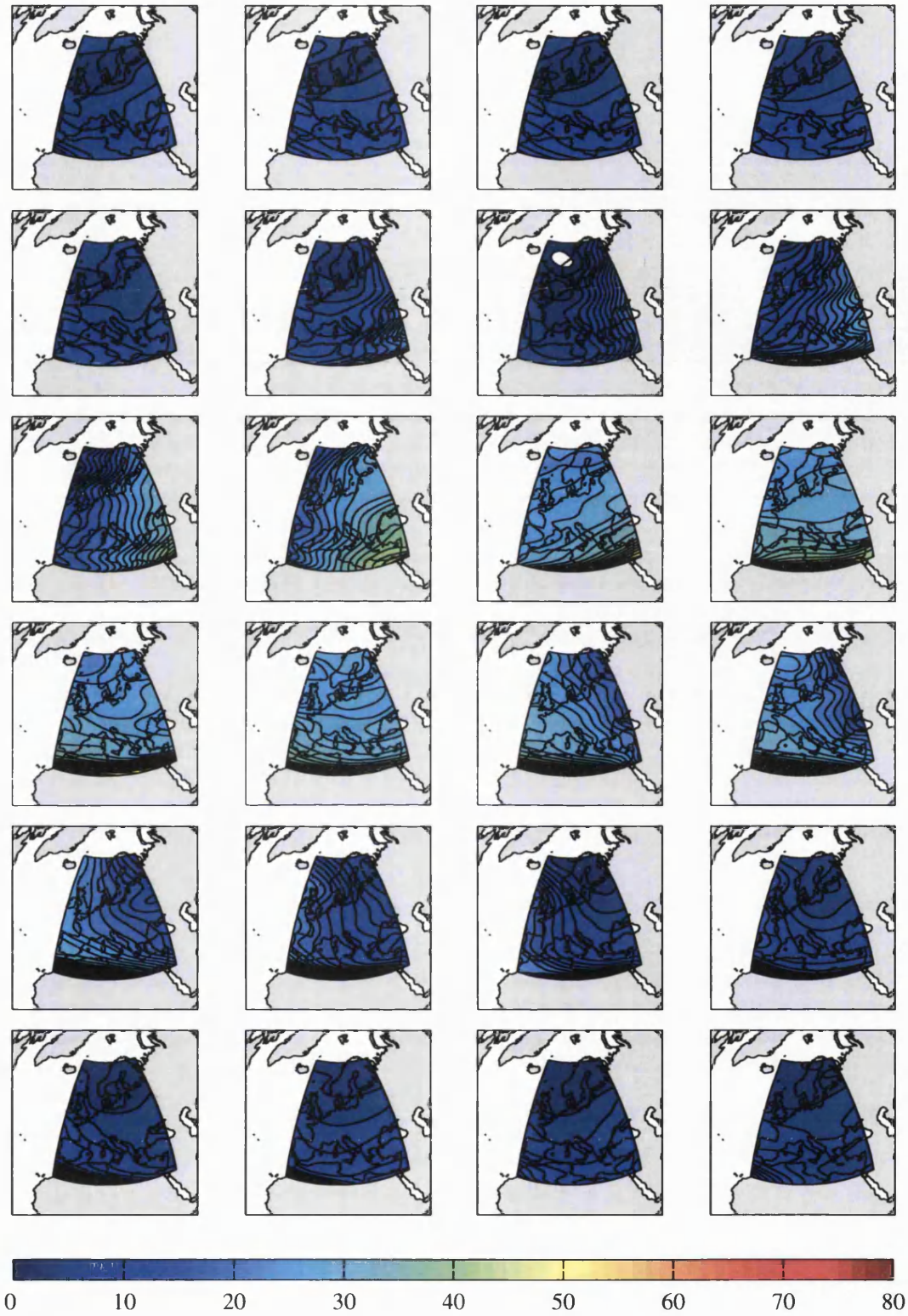


Figure A.3: Hourly vertical TEC maps for 04 February 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

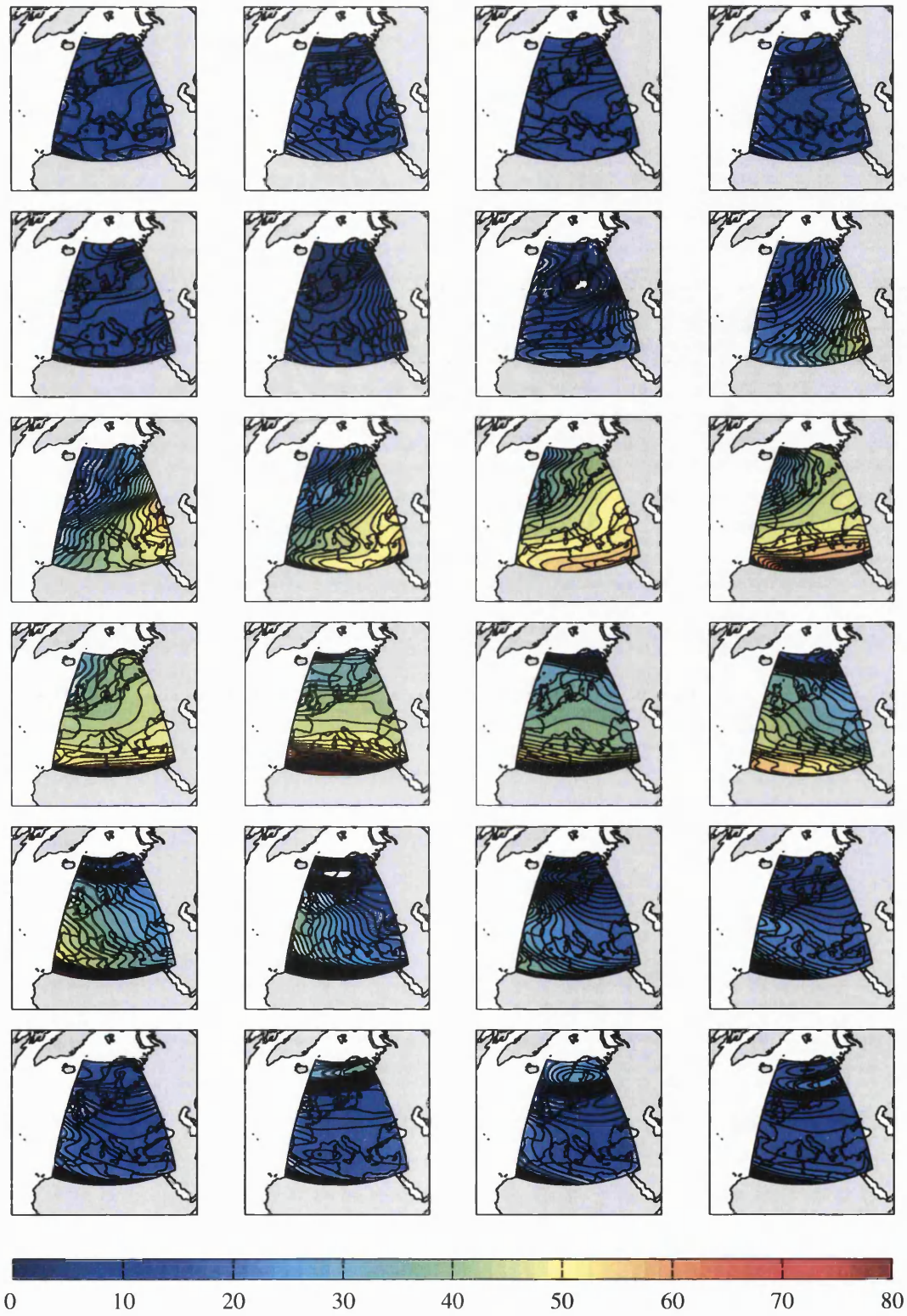


Figure A.4: Hourly vertical TEC maps for 13 February 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

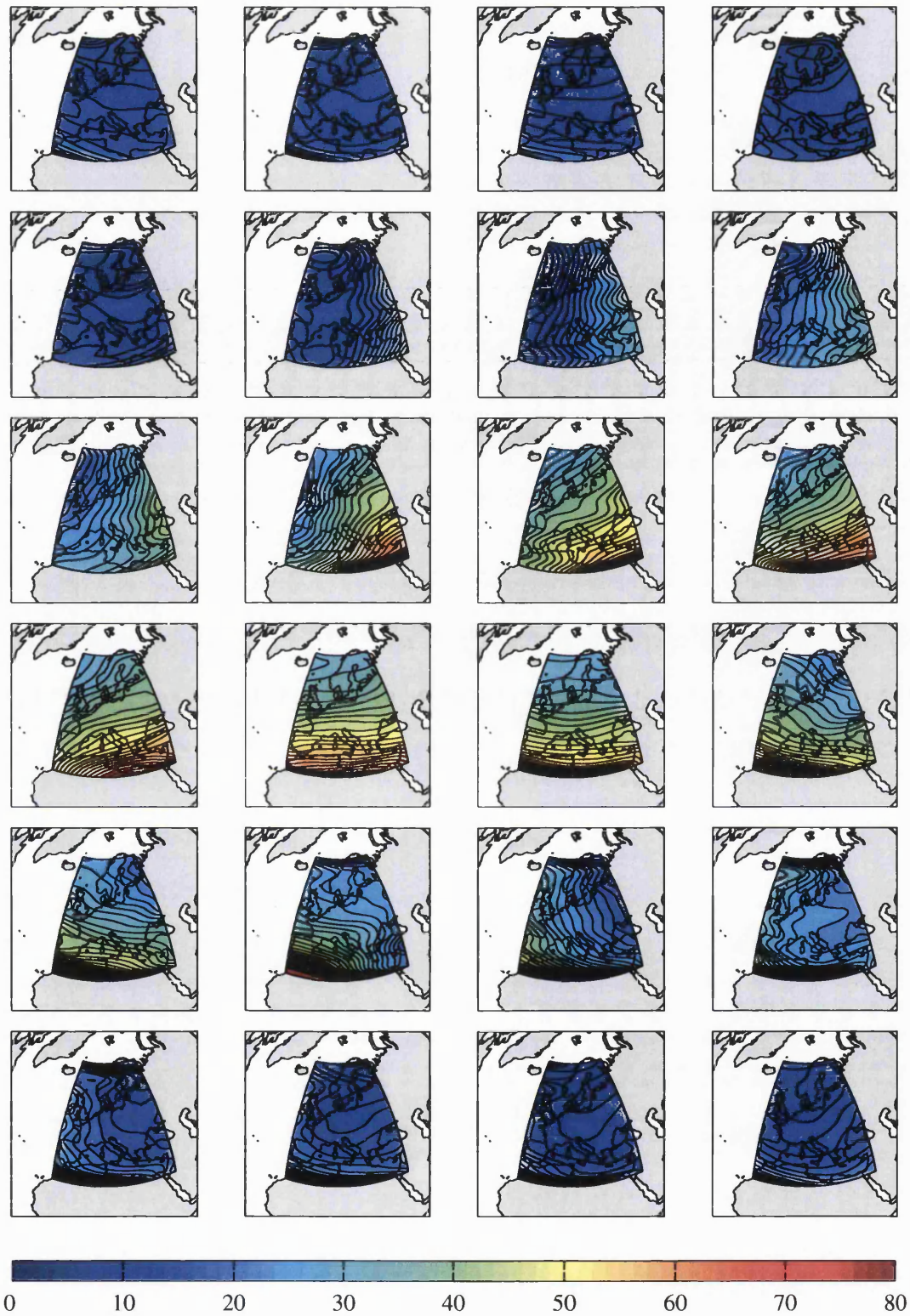


Figure A.5: Hourly vertical TEC maps for 16 March 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

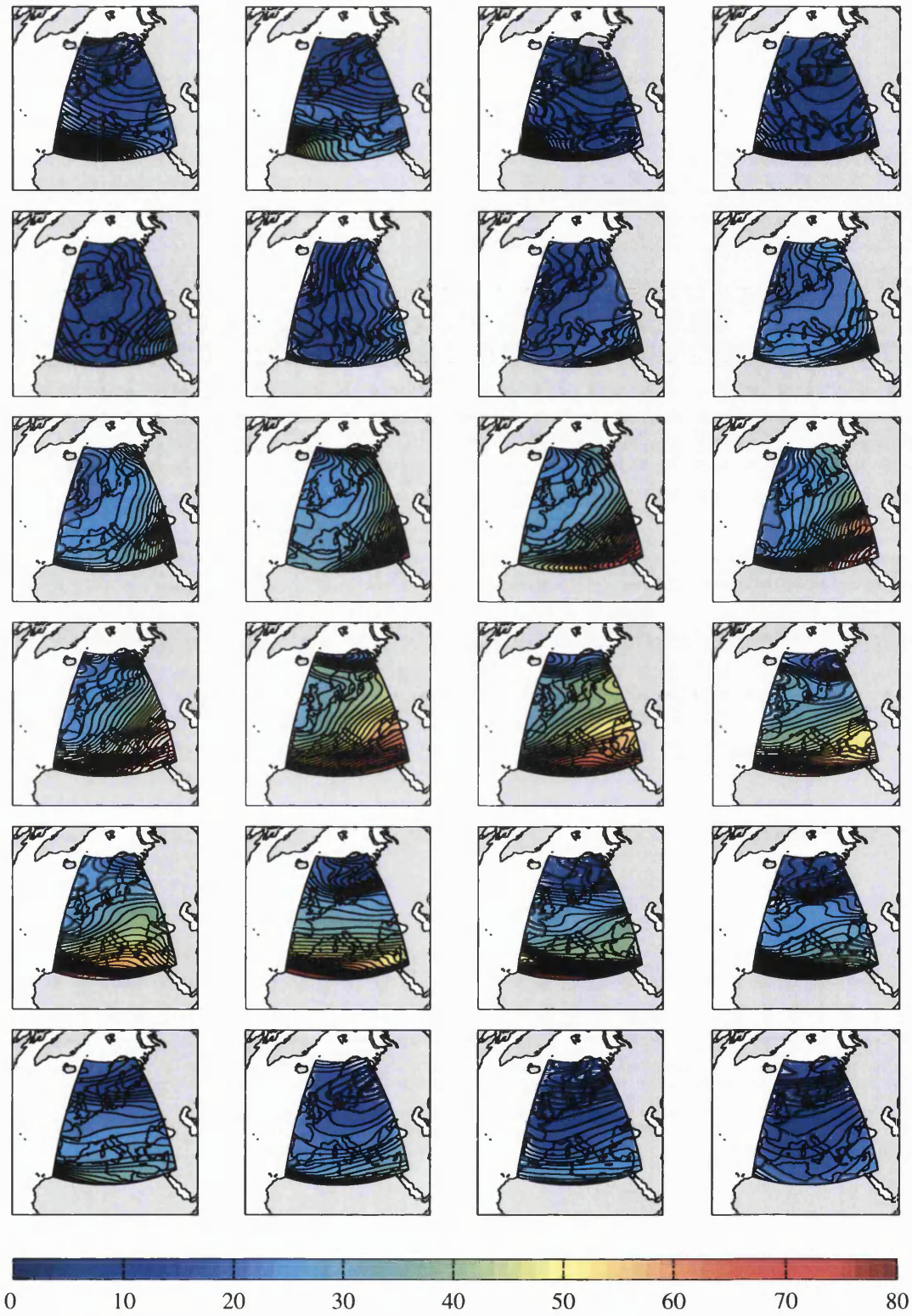


Figure A.6: Hourly vertical TEC maps for 28 March 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

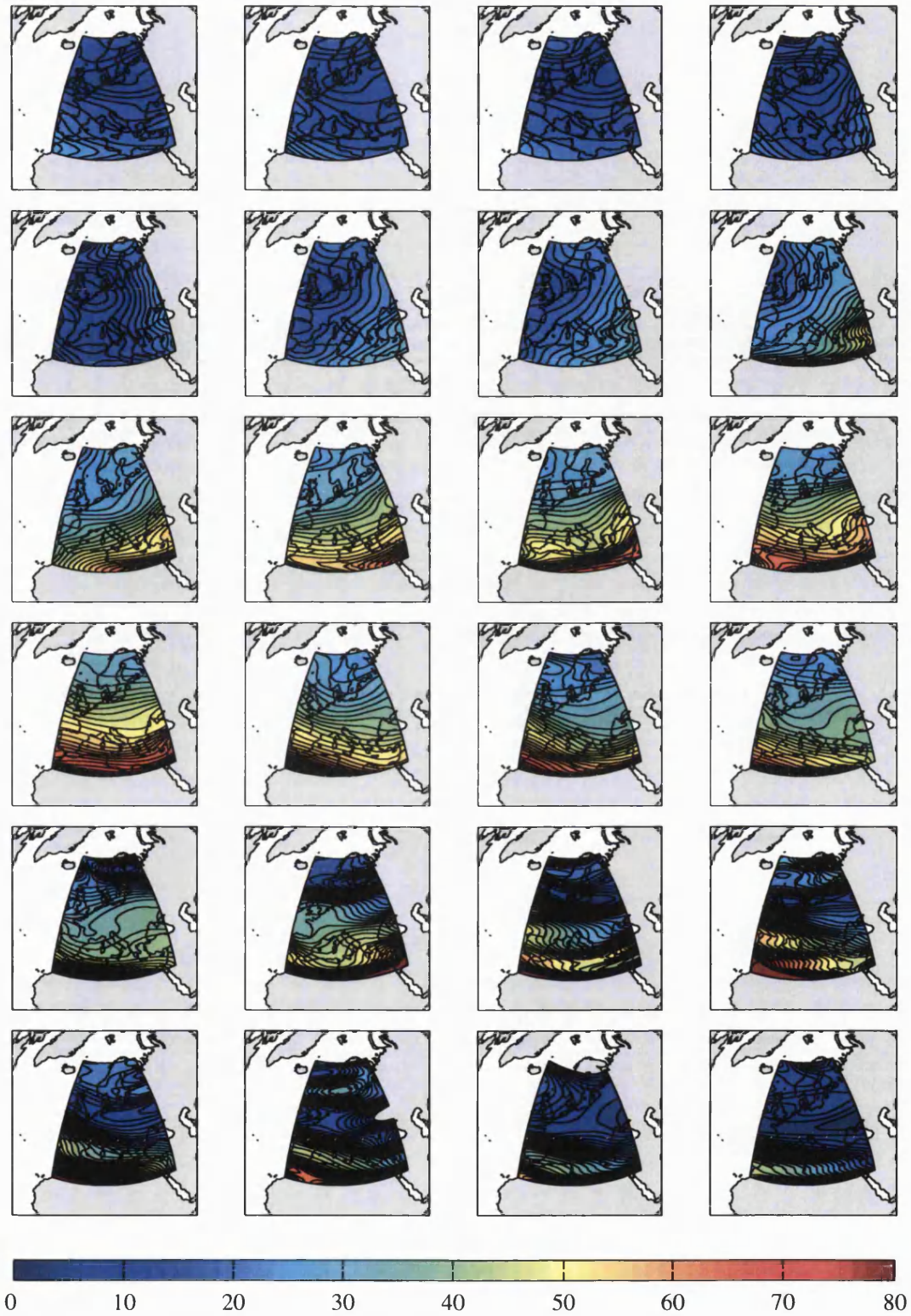


Figure A.7: Hourly vertical TEC maps for 11 April 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

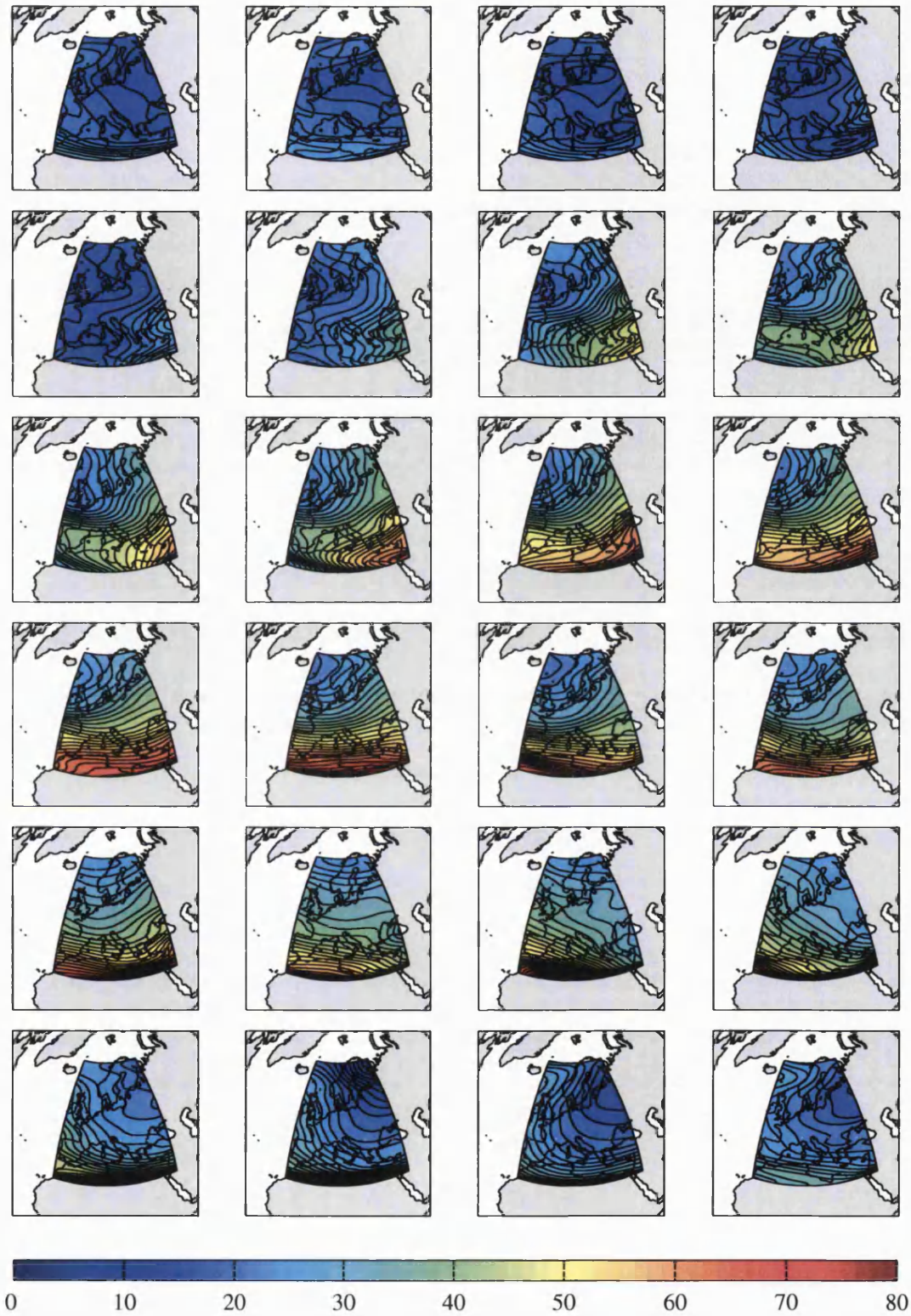


Figure A.8: Hourly vertical TEC maps for 30 April 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

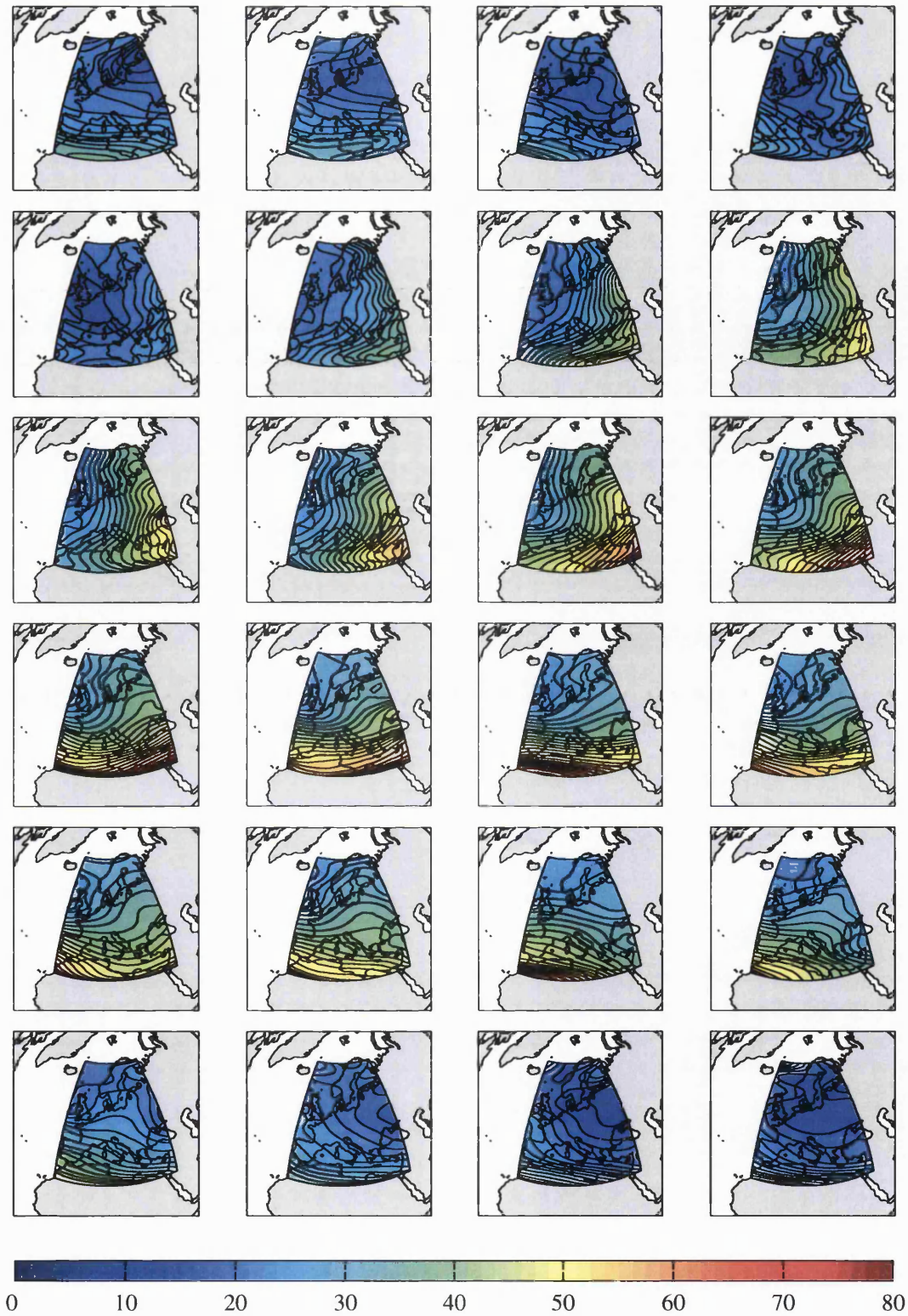


Figure A.9: Hourly vertical TEC maps for 06 May 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

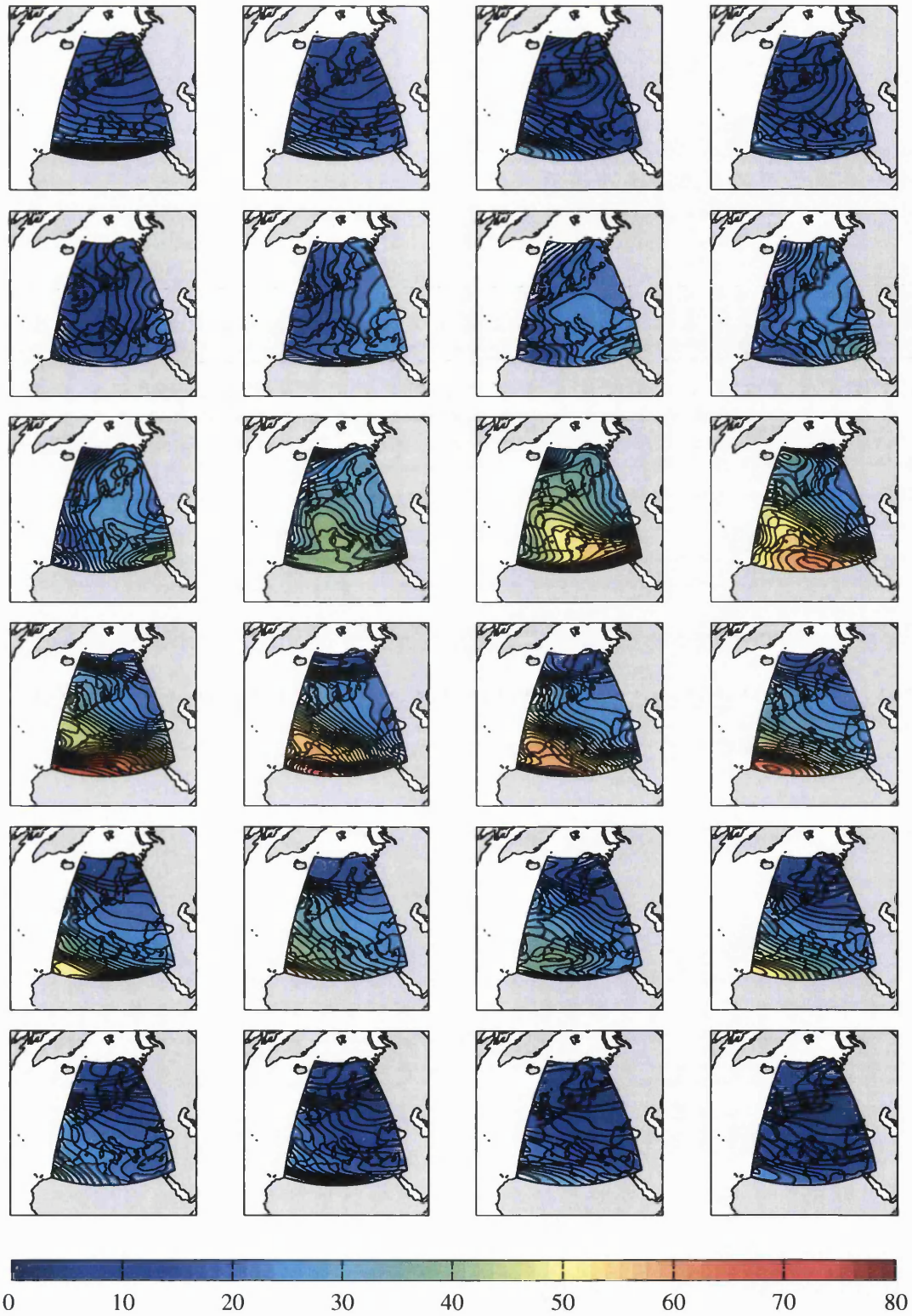


Figure A.10: Hourly vertical TEC maps for 12 May 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

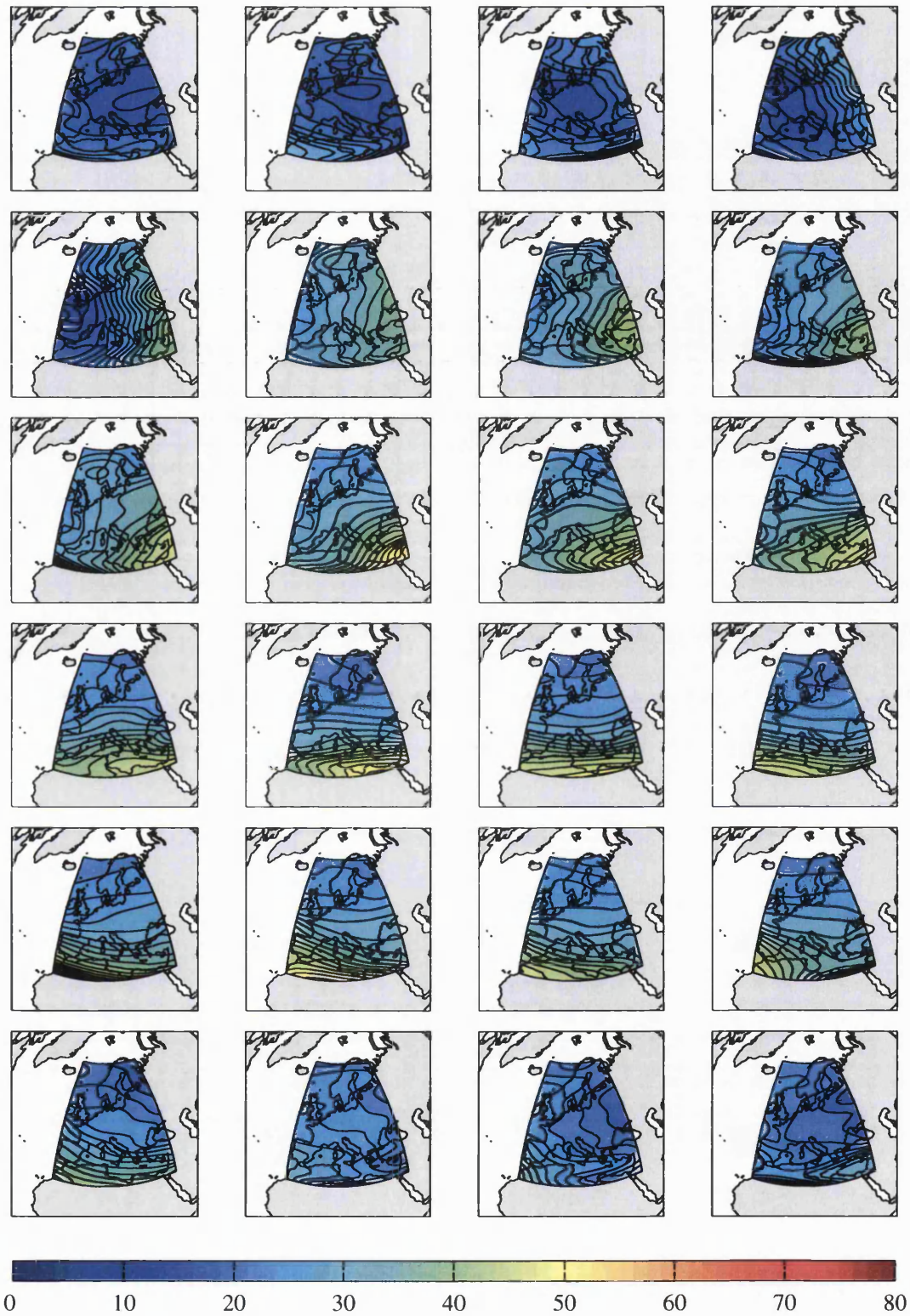


Figure A.11: Hourly vertical TEC maps for 31 May 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

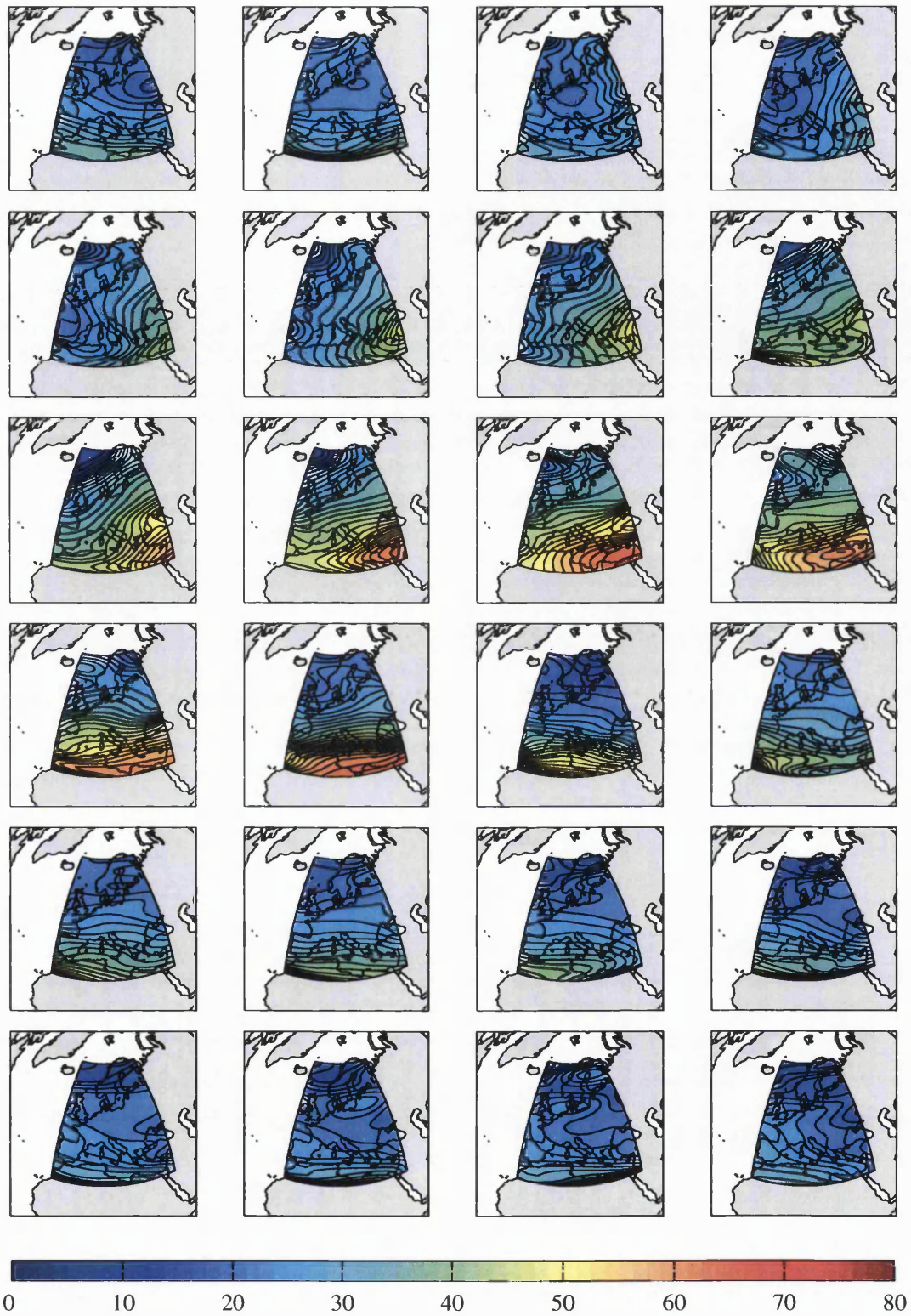


Figure A.12: Hourly vertical TEC maps for 18 June 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

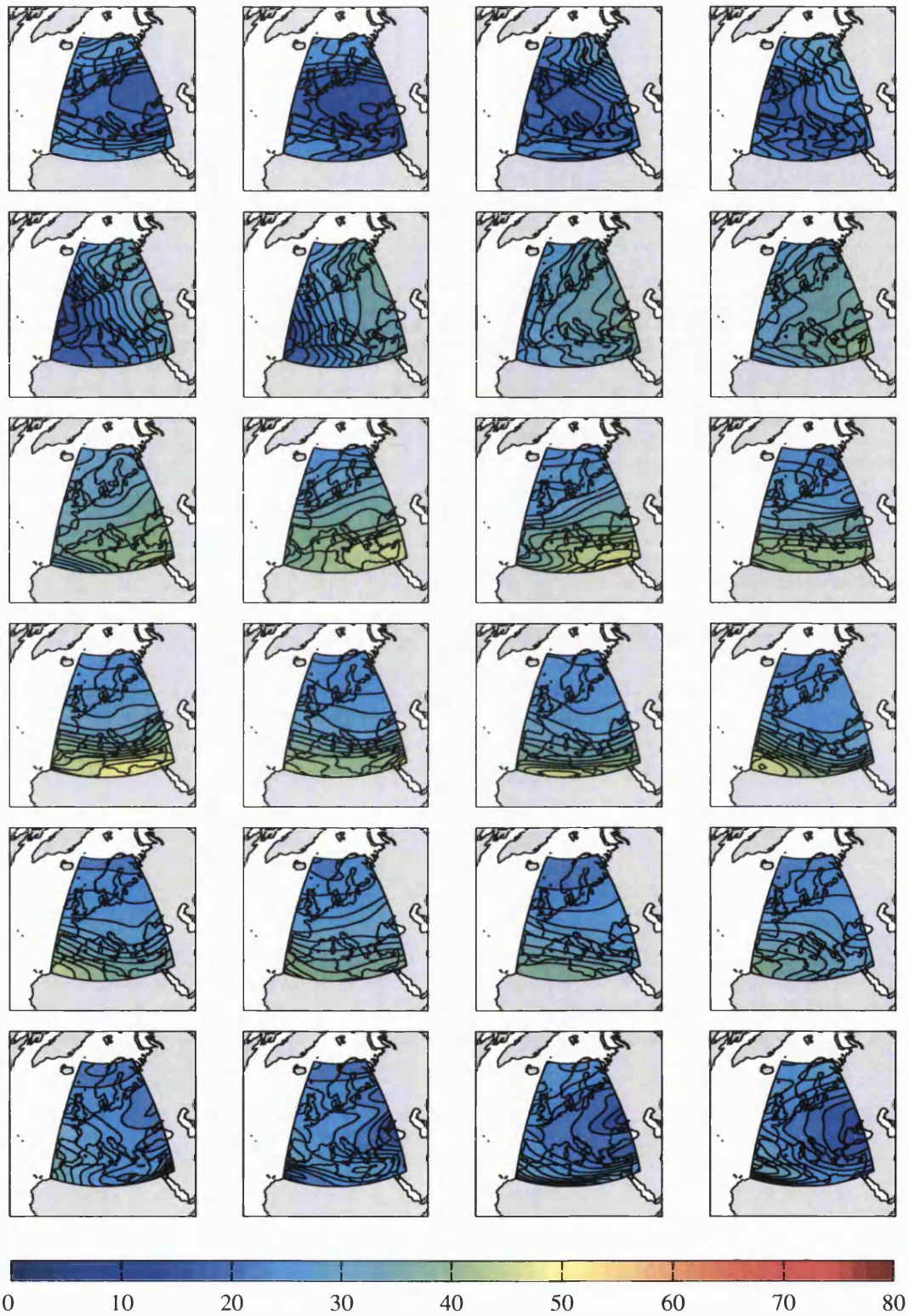


Figure A.13: Hourly vertical TEC maps for 28 June 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

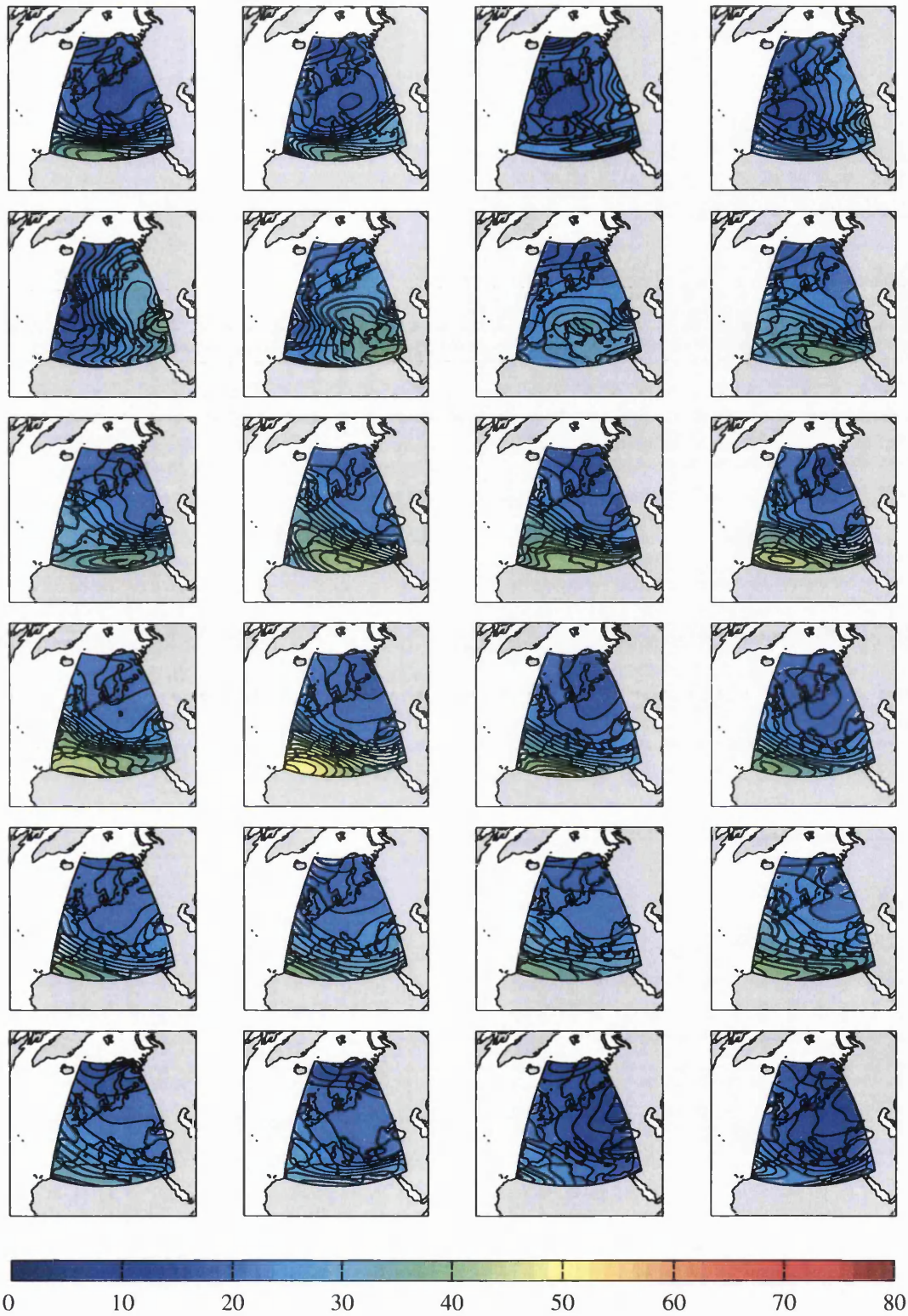


Figure A.14: Hourly vertical TEC maps for 04 July 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

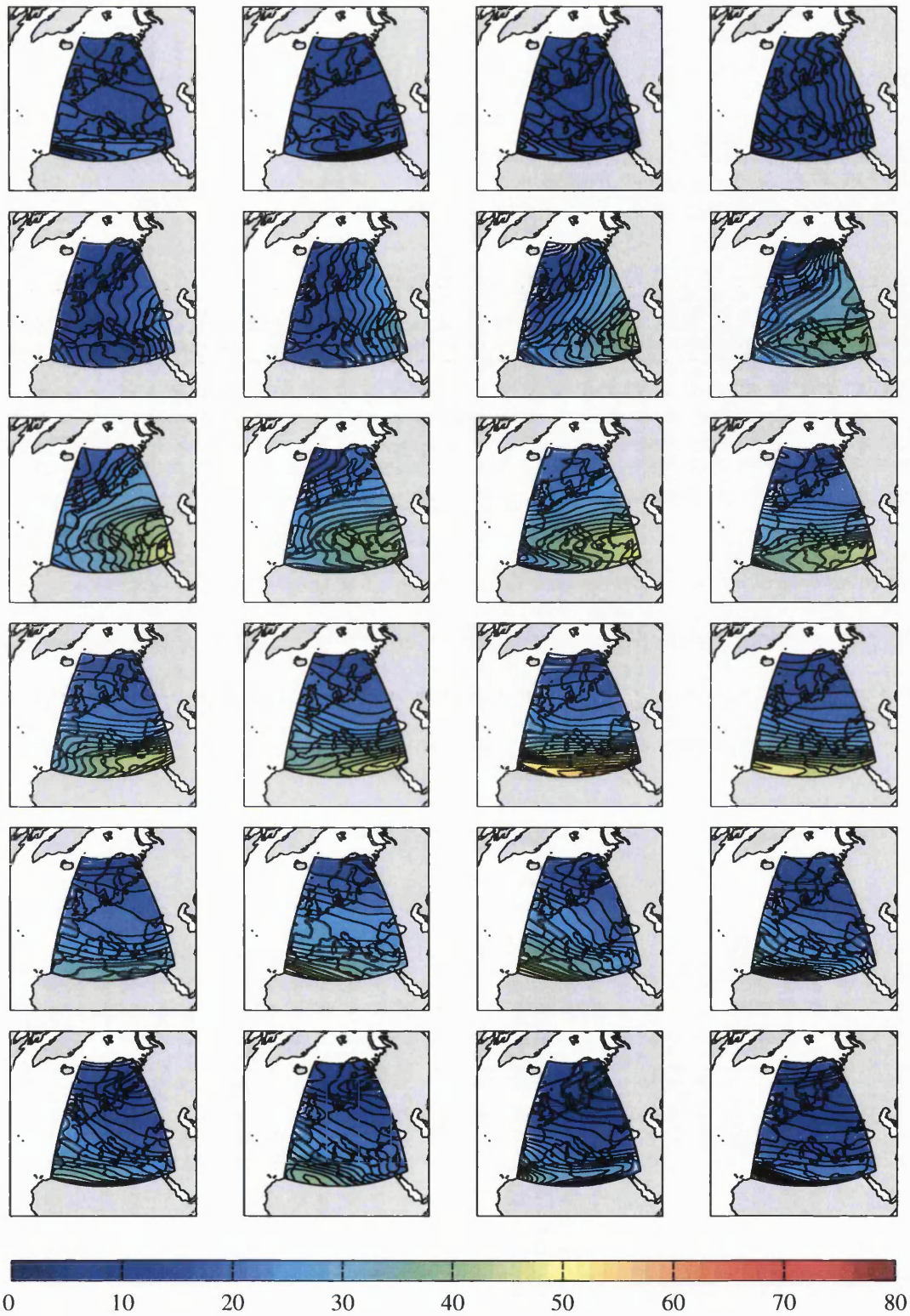


Figure A.15: Hourly vertical TEC maps for 25 July 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

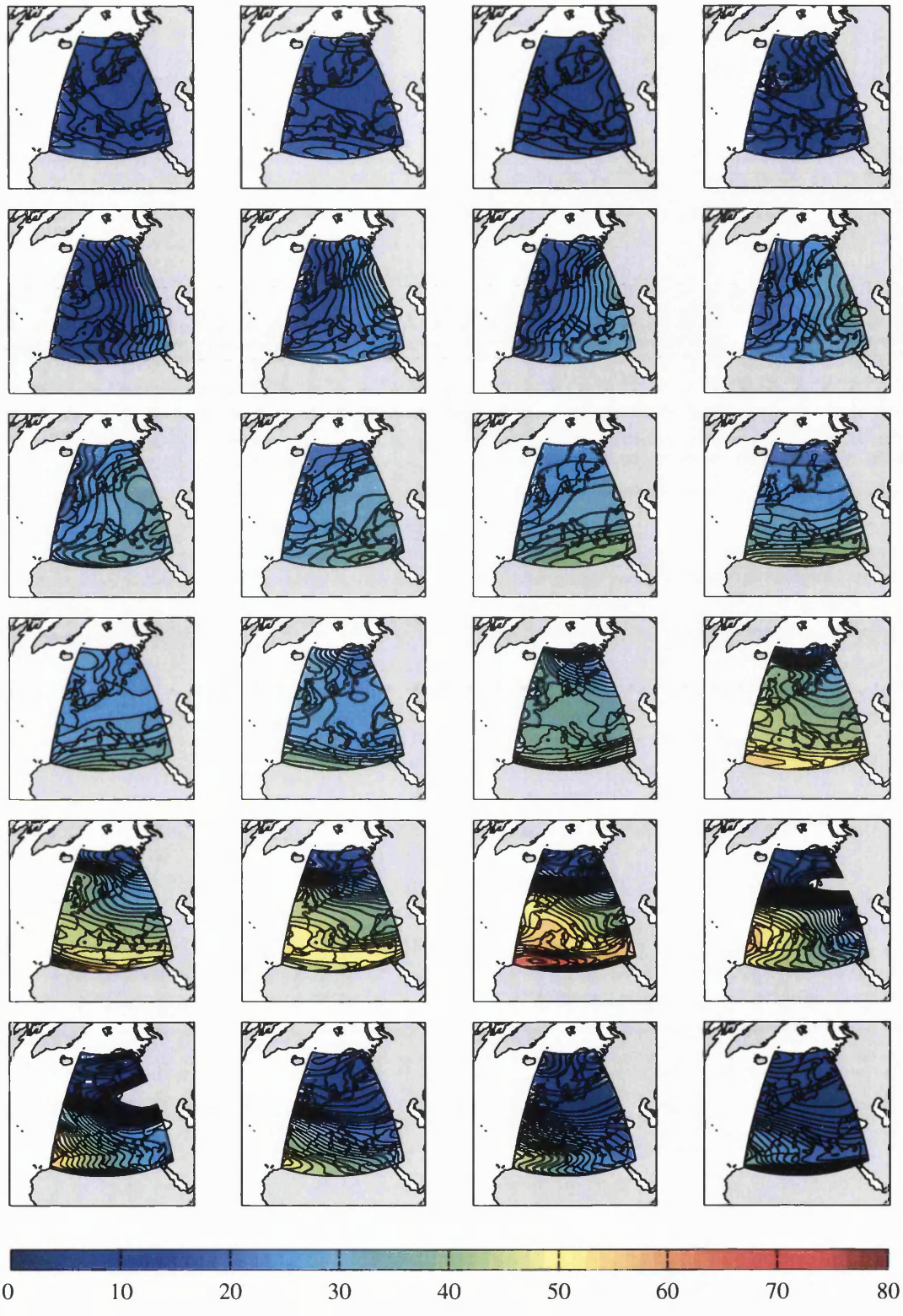


Figure A.16: Hourly vertical TEC maps for 17 August 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

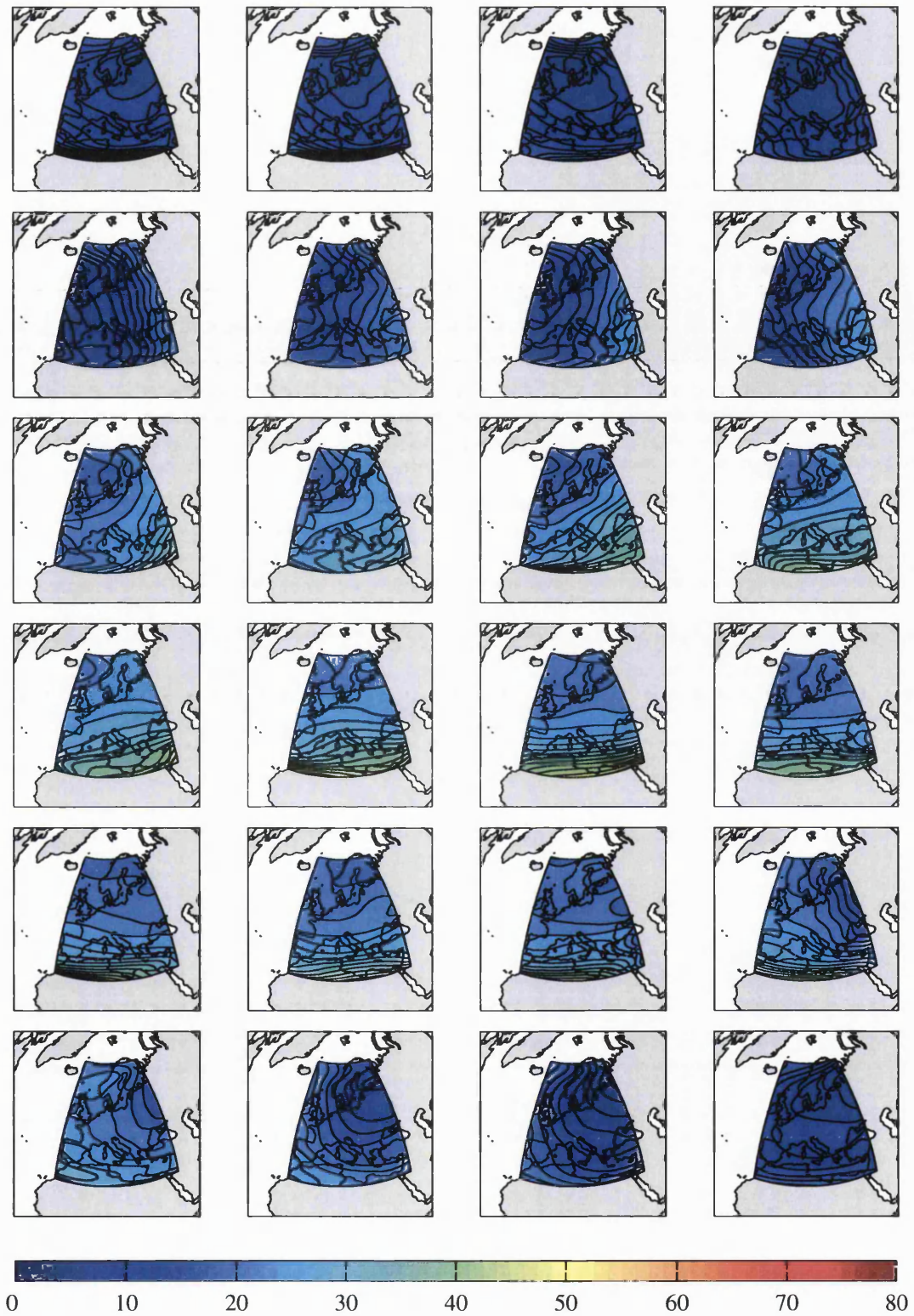


Figure A.17: Hourly vertical TEC maps for 24 August 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

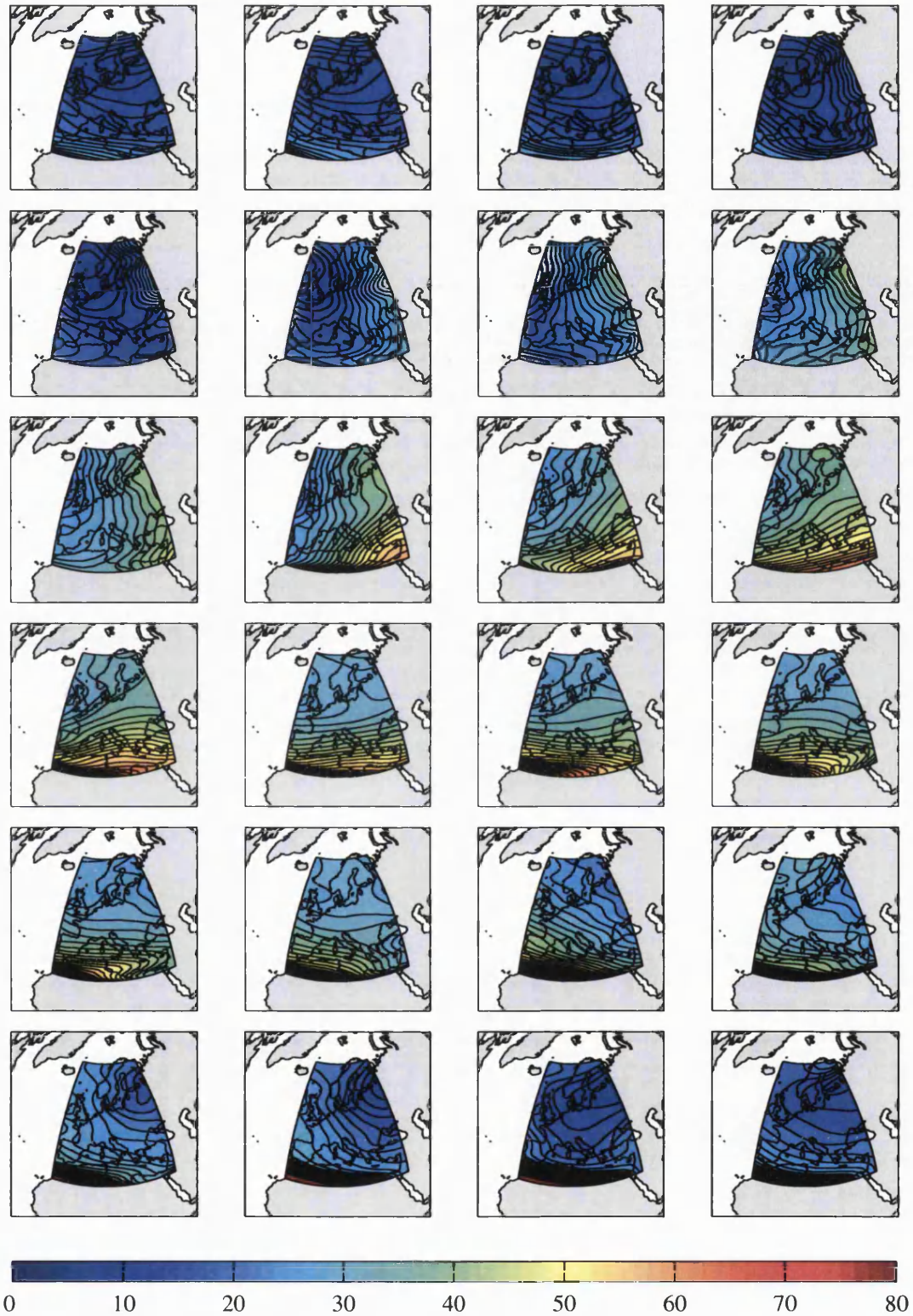


Figure A.18: Hourly vertical TEC maps for 07 September 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

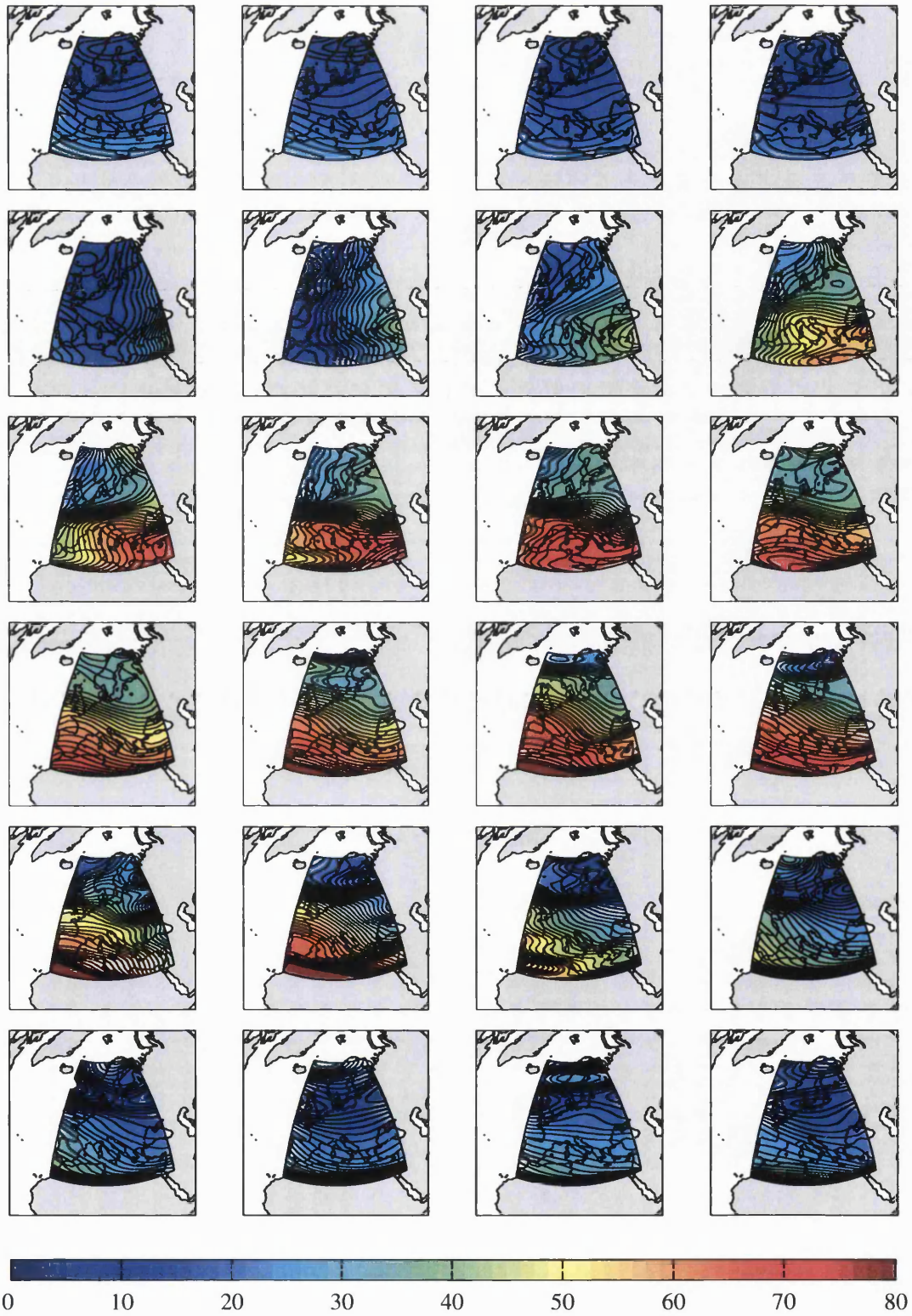


Figure A.19: Hourly vertical TEC maps for 23 September 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

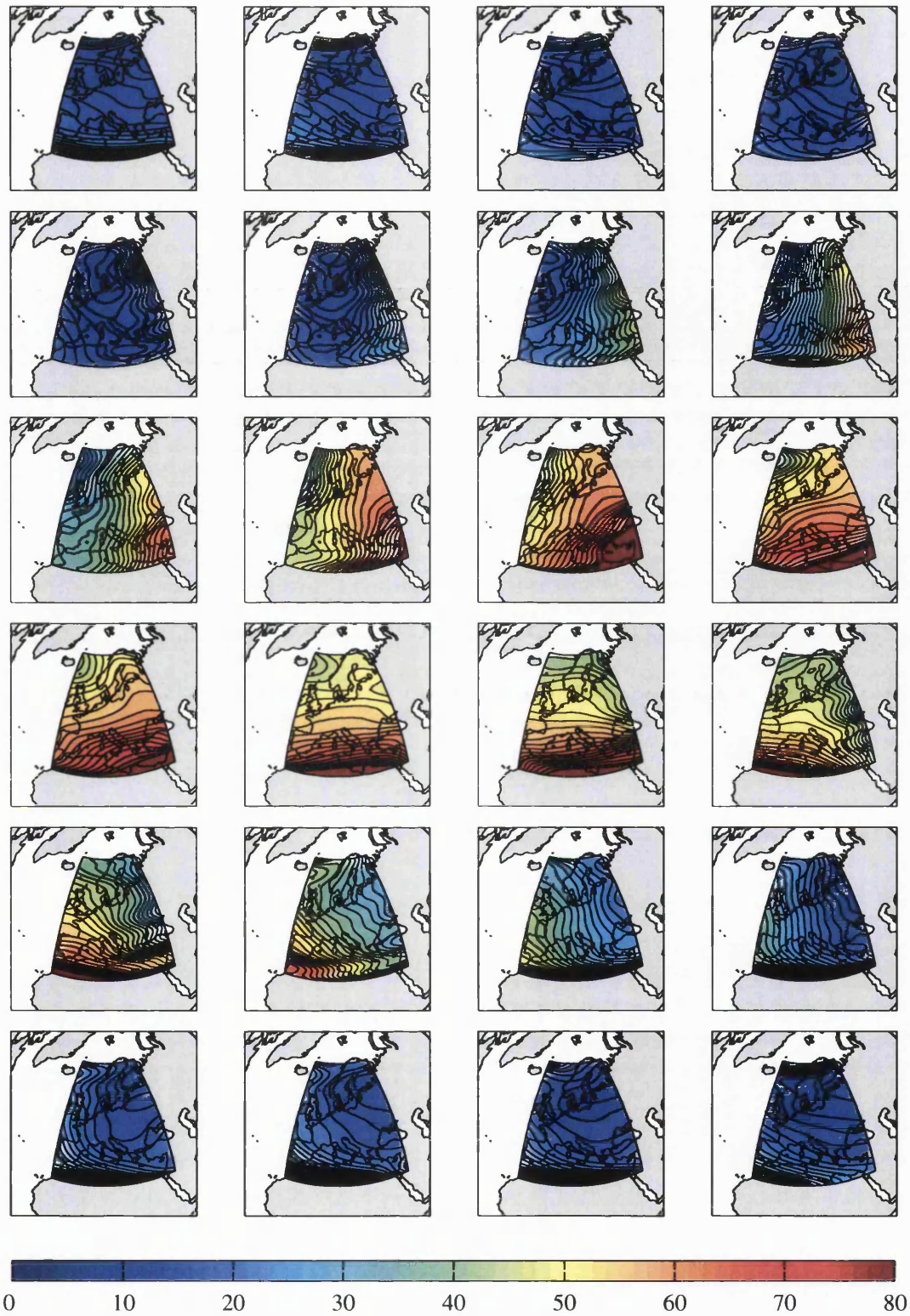


Figure A.20: Hourly vertical TEC maps for 07 October 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

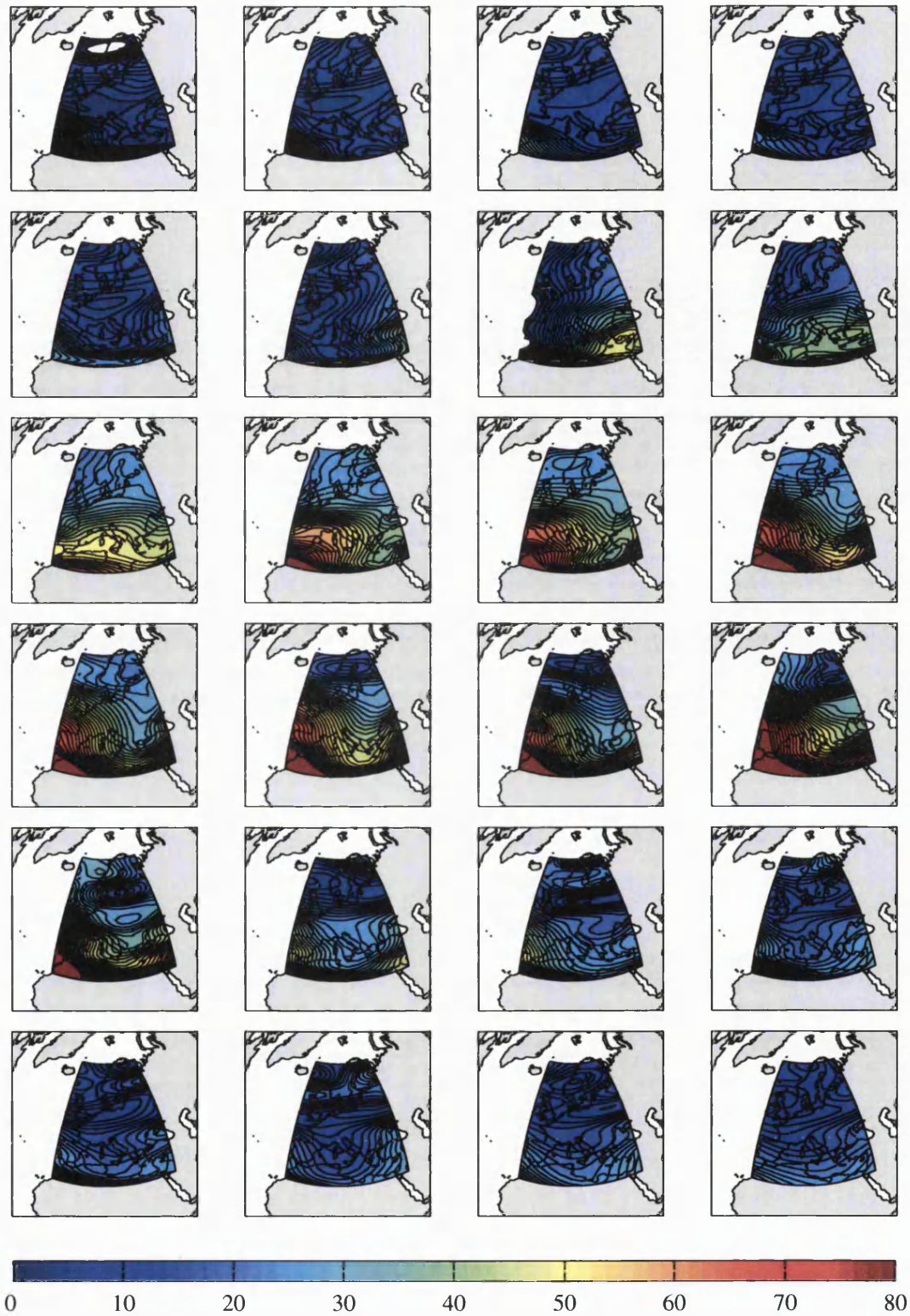


Figure A.21: Hourly vertical TEC maps for 22 October 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

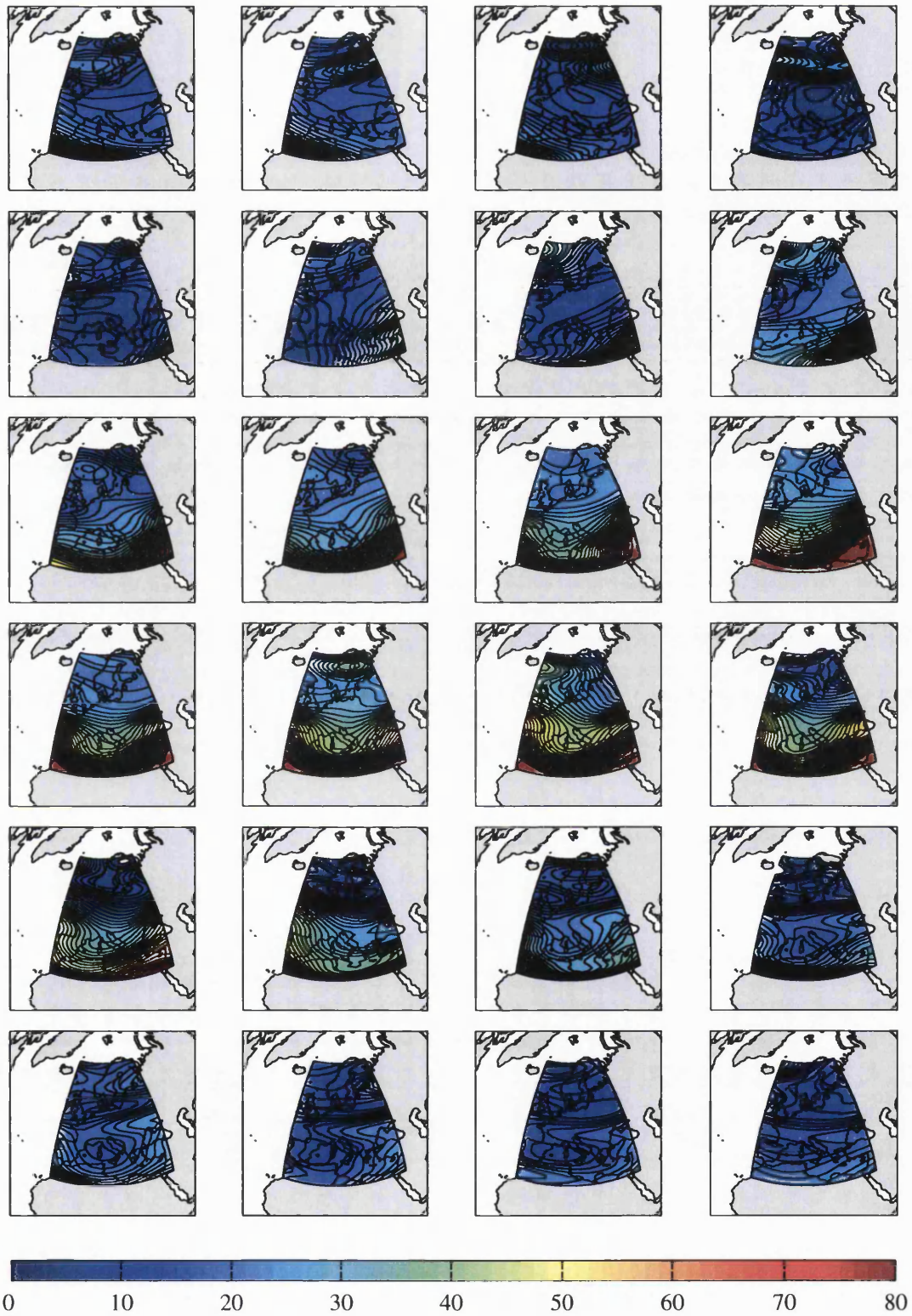


Figure A.22: Hourly vertical TEC maps for 06 November 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

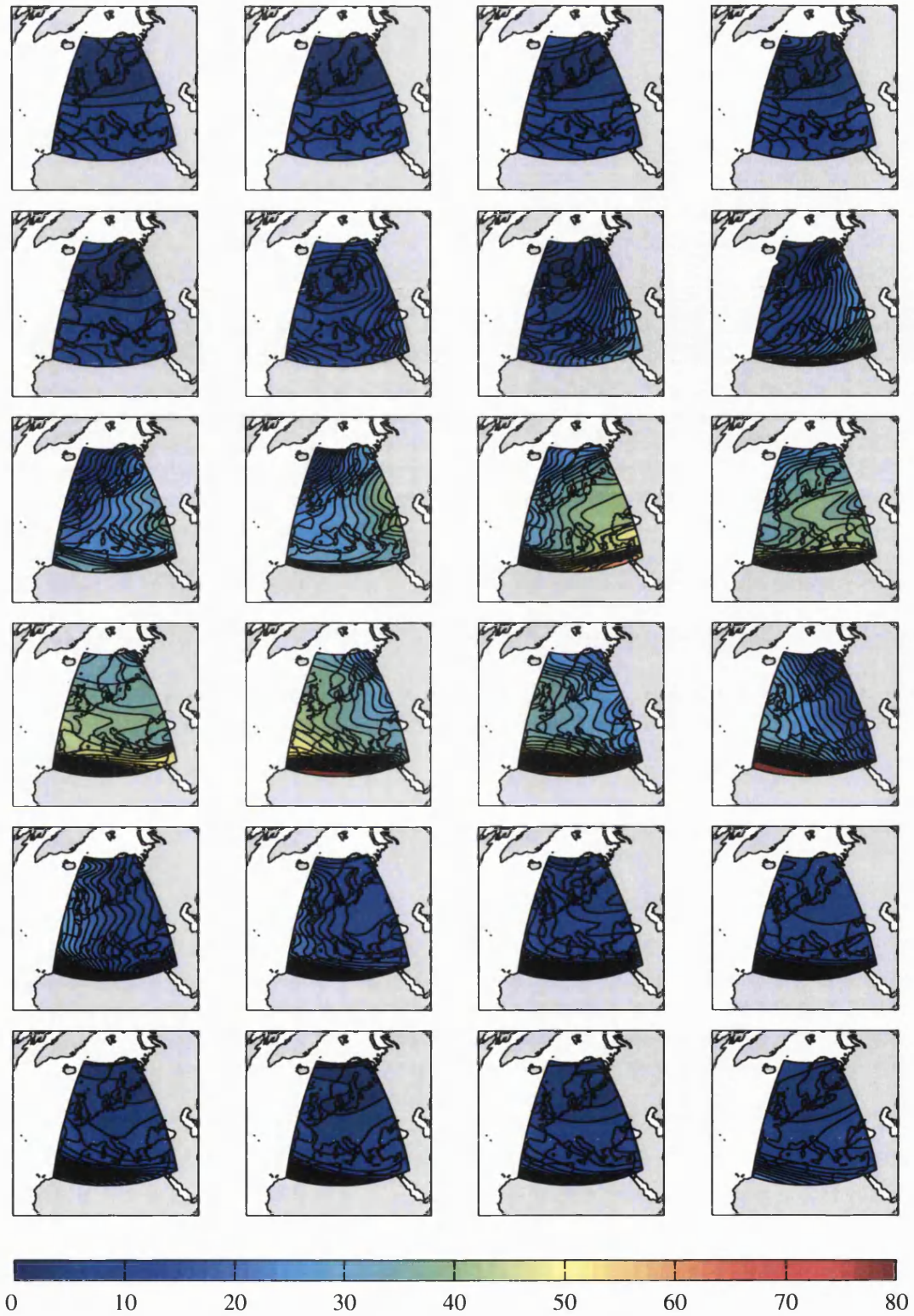


Figure A.23: Hourly vertical TEC maps for 27 November 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

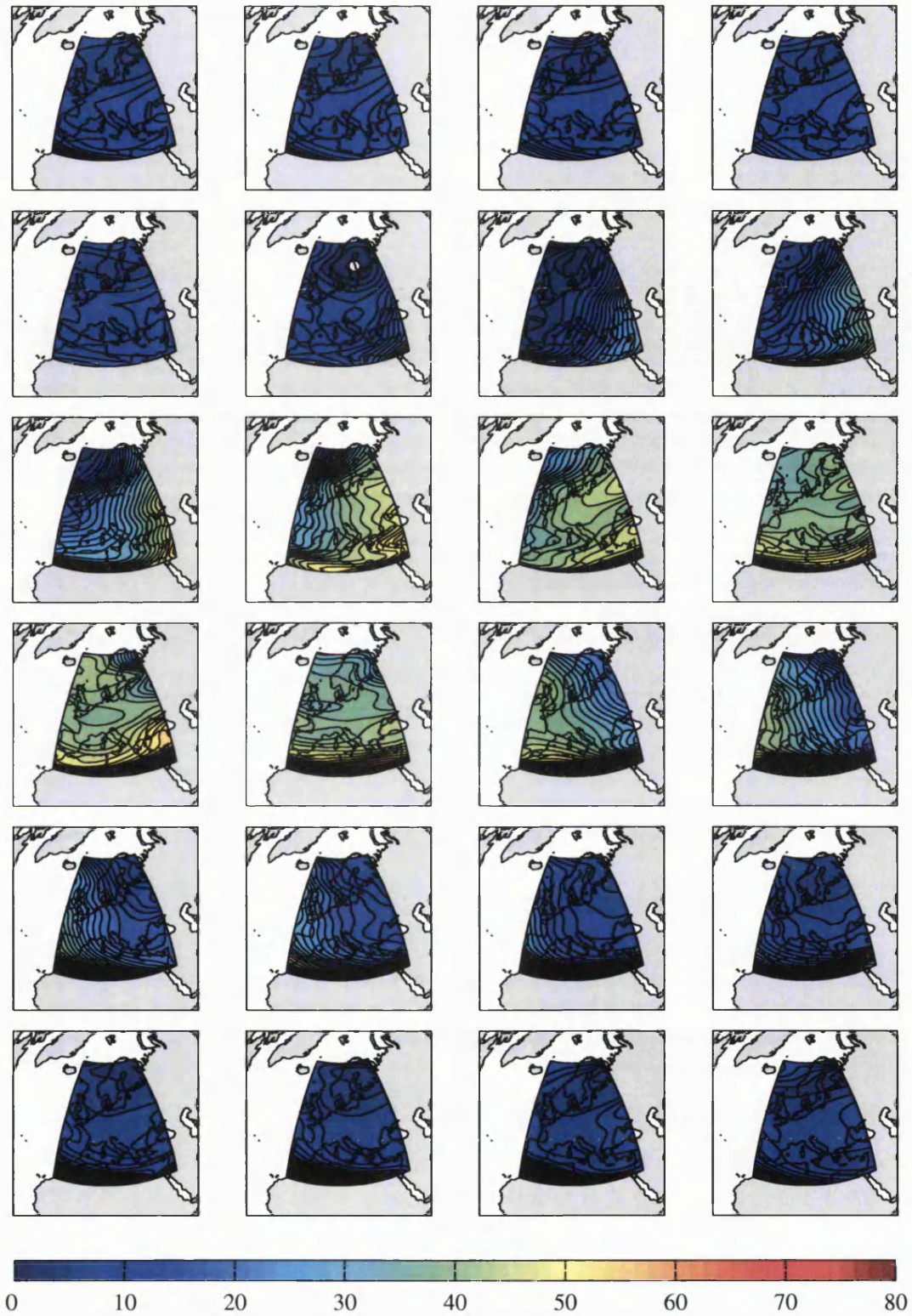


Figure A.24: Hourly vertical TEC maps for 09 December 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

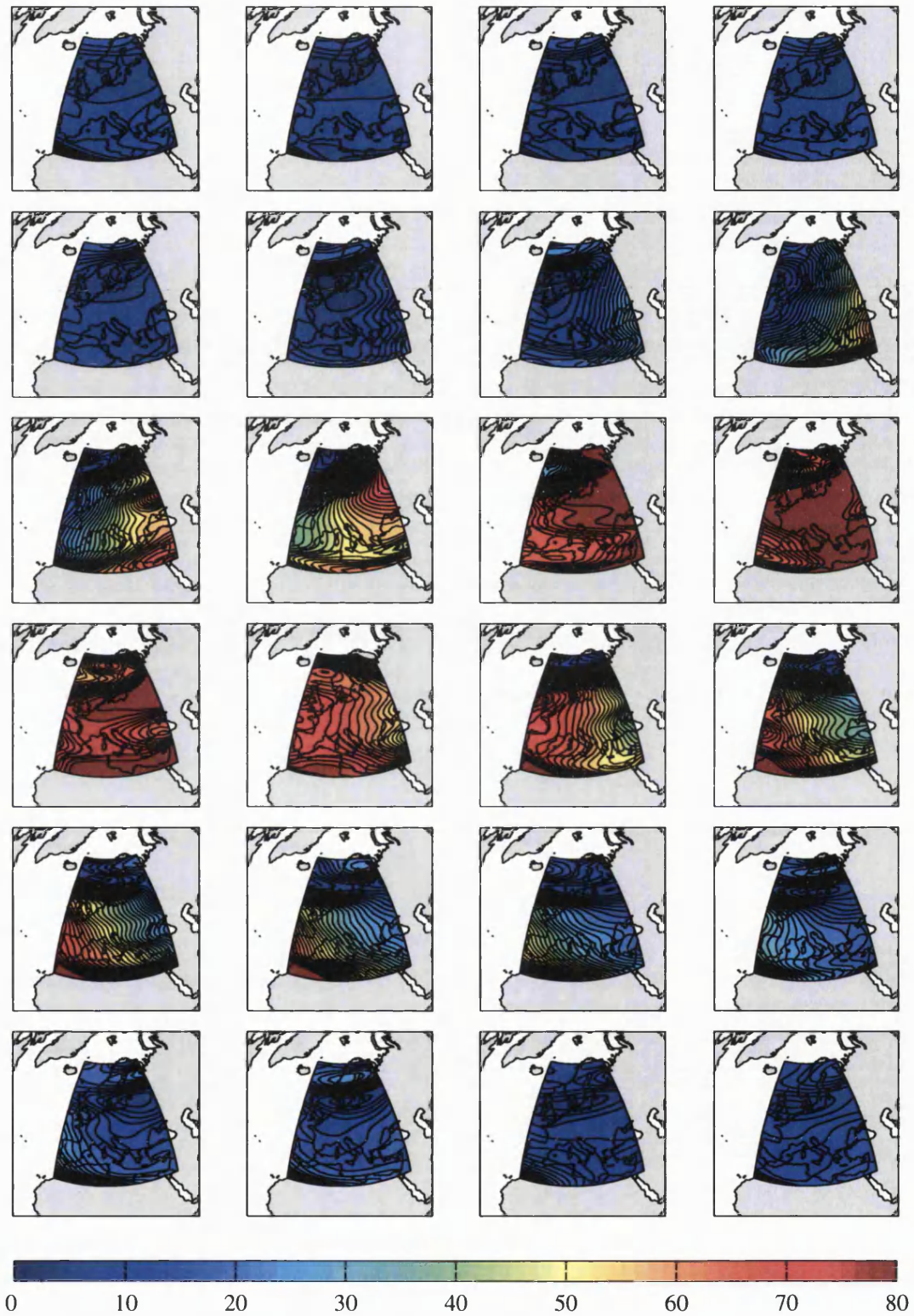


Figure A.25: Hourly vertical TEC maps for 24 December 2001. Hours progress from 00 UT at top left to 23 UT at bottom right.

Appendix B

Publications

B.1 Journal papers

1. **Meggs R. W., Mitchell C. N. and Howells V. S. C.**. "Simultaneous observations of the main trough using GPS imaging and the EISCAT radar", *Annales Geophysicae*, in press.

2. **Meggs R. W., Mitchell C. N. and Spencer, P. S. J.** "A comparison of techniques for mapping total electron content over Europe using GPS signals", *Radio Science*, 39, RS1S10, doi:10.1029/2002RS002846, Feb. 2004.

B.2 Conference proceedings

1. **Robert W. Meggs Cathryn N. Mitchell and Glyn Wyman**, "Use of GPS for mapping Total Electron Content over Europe", *Proceedings of the Beacon Satellite Symposium (BSS 2004)*, The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy October 18 - 22, 2004.

2. **R. W. Meggs, C. N. Mitchell and V. S. C. Howells**, "Simultaneous observations of the main trough using GPS imaging and the EISCAT radar", *Special Symposium of the URSI Joint Working Group FG on Atmospheric Remote Sensing using Satellite Navigation Systems*, ASI Centro di Geodesia Spaziale "Giuseppe Colombo", Matera, Italy, 13 - 15 October 2003.

3. **Richard M. Dear, Robert W. Meggs, Cathryn N. Mitchell**"Determination of GPS satellite and receiver interfrequency biases", *Special Symposium of the URSI Joint Working Group FG on Atmospheric Remote Sensing using Satel-*

lite Navigation Systems, ASI Centro di Geodesia Spaziale "Giuseppe Colombo", Matera, Italy, 13 - 15 October 2003.

4. **Robert W Meggs, Cathryn N Mitchell and Paul S J Spencer**, "Simulations of thin shell and 4-D inversion techniques for mapping of total electron content", *27th General Assembly of the Union of Radio Science International (URSI)*, Maastricht, Netherlands, 17 - 24 August 2002.

5. **Meggs R. W., Mitchell C. N. and Spencer, P. S. J.** "Comparison of Thin Shell and 4-D Inversion Techniques for Measurement and Mapping of Total Electron Content Using GPS Signals Over Europe." *Proceedings of the 10th Ionospheric Effects Symposium, Alexandria*, VA, USA, 7 - 9 May 2002.

6. **Cathryn N Mitchell, Robert W. Meggs and Paul S J Spencer** "4-D inversion for measurement and mapping of total electron content using GNSS signals", Meeting of the Satellite-Based Augmentation System - Ionosphere group, at ESA-ESTEC, Netherlands, May 2001.